Theoretical Physics I-Mathematical Methods

Daniel Erenso and Victor Montemayor

April 15, 2019
Contents

Introduction xi

I Mathematical Methods in Physics I 1

1 Series and Convergence 3

1.1 Sequence and series . . . . . . . . . . . . . . . . . . . . . . . . . . 3
1.2 Testing series for convergence . . . . . . . . . . . . . . . . . . . . 7
1.3 Series representations of real functions . . . . . . . . . . . . . . 9
1.4 Homework Assignment 1 . . . . . . . . . . . . . . . . . . . . . . . 18
1.5 Homework Assignment 2 . . . . . . . . . . . . . . . . . . . . . . . 19

2 Complex Numbers, functions, and series 21

2.1 Complex numbers . . . . . . . . . . . . . . . . . . . . . . . . . . 21
2.2 Complex Infinite Series and the Circle of Convergence . . . . . . 28
2.3 Powers and Roots of complex numbers . . . . . . . . . . . . . . 32
2.4 Algebraic vs transcendental functions . . . . . . . . . . . . . . . 34
2.5 Homework Assignment 3 . . . . . . . . . . . . . . . . . . . . . . . 40
2.6 Homework Assignment 4 . . . . . . . . . . . . . . . . . . . . . . . 42
2.7 Homework Assignment 5 . . . . . . . . . . . . . . . . . . . . . . . 43

3 Vectors, Lines, and Planes 45

3.1 An Overview of Vector Fundamentals . . . . . . . . . . . . . . . 45
3.2 Vector Multiplication . . . . . . . . . . . . . . . . . . . . . . . . . 46
3.3 Equations of a straight line . . . . . . . . . . . . . . . . . . . . . 49
3.4 Equation of a plane . . . . . . . . . . . . . . . . . . . . . . . . . . 49
3.5 Homework Assignment 6 . . . . . . . . . . . . . . . . . . . . . . . 53

4 Matrix and determinants 55

4.1 Important Terminologies . . . . . . . . . . . . . . . . . . . . . . . 55
4.2 Matrix Arithmetic and Manipulation . . . . . . . . . . . . . . . 56
4.3 Matrix representation of a set of linear equations . . . . . . . . 57
4.4 Matrix application: solving a set of linear equations . . . . . . 58
4.5 Determinant of a square matrix . . . . . . . . . . . . . . . . . . . 63
CONTENTS

4.6 Cramer’s Rule .......................... 66
4.7 The adjoint and inverse of a matrix .................... 67
4.8 Orthogonal matrices and the rotation matrix .......... 72
4.9 Linear independence .......................... 74
4.10 Gram-Schmidt orthogonalization ....................... 78
4.11 Homework Assignment 7 .......................... 80
4.12 Homework Assignment 8 .......................... 81
4.13 Homework Assignment 9 .......................... 83

5 Introduction to differential calculus-I 85
5.1 Partial differentiation .......................... 85
5.2 Total differential .............................. 89
5.3 The multivariable form of the Chain rule ............ 90
5.4 Extremum (Max/Min) problems ..................... 94
5.5 The Method of Lagrangian Multipliers ............... 97
5.6 Change of Variables ............................ 101
5.7 Legendre Transformations ........................ 106
5.8 Homework Assignment 10 .......................... 107
5.9 Homework Assignment 11 .......................... 108

6 Introduction to Differential Calculus II 109
6.1 Linear ordinary differential equations ................. 109
6.1.1 First-order LODE with constant coefficients .... 110
6.2 The first-order DE and exact differential ............. 117
6.3 First-order DE and none exact total differential ....... 122
6.4 Higher-Order ODE’s ............................ 126
6.4.1 Homogeneous ODE’s with constant coefficient .. 127
6.4.2 None homogeneous ODE’s with Constant Coefficients . 137
6.5 The method of superposition and the particular solution. 141
6.6 The method of successive integration ................ 145
6.7 Partial Differential Equations ........................ 154
6.8 Homework Assignment 12 .......................... 159
6.9 Homework Assignment 13 .......................... 159
6.10 Homework Assignment 14 .......................... 160
6.11 Homework Assignment 15 .......................... 161

7 Integral calculus-scalar functions 163
7.1 Integration in Cartesian coordinates .................. 163
7.1.1 Single and multiple integrals .................... 163
7.1.2 Physical Applications .......................... 171
7.2 Integration in curvilinear coordinates ................ 183
7.2.1 One- and two-dimensional curvilinear Coordinates .... 184
7.2.2 3-D Curvilinear coordinates-cylindrical ............ 186
7.2.3 3-D Curvilinear Coordinate Systems: spherical Coordinates 188
7.3 Homework Assignment 16 .......................... 191
7.4 Homework Assignment 17 .......................... 192
CONTENTS

7.5 Homework Assignment 18 ........................................ 193

8 Vector Calculus .................................................. 195
  8.1 Vector Products .............................................. 195
     8.1.1 Physical applications .................................. 196
     8.1.2 Vectors derivatives .................................... 198
  8.2 Differential Operators ...................................... 202
     8.2.1 The gradient operator and directional derivative .... 202
     8.2.2 The Divergence, the Curl, and the Laplacian ......... 206
  8.3 Single vector function integrals ............................ 217
     8.3.1 Line vector integrals .................................. 217
     8.3.2 Conservative vector fields ............................ 219
     8.3.3 Conservative Fields and Exact Differentials ......... 221
  8.4 Multiple vector integrals ................................... 222
     8.4.1 Green’s Theorem ...................................... 222
     8.4.2 The Stokes’ Theorem .................................. 226
     8.4.3 The Divergence Theorem ............................... 231
     8.4.4 More Examples on Divergence and Stokes theorems .. 233
     8.4.5 Applications of the Divergence and Stoke’s theorems . 238

II Mathematical Methods in Physics II ............................ 241

9 Introduction to the Calculus of Variations ..................... 243
  9.1 Geodesic and stationary points ............................. 243
  9.2 The general problem ...................................... 247
  9.3 Applications: The Brachystochrone Problem ............... 250
  9.4 Applications: classical mechanics ......................... 253
  9.5 Physical application of the Euler-Lagrange equation ...... 254

10 Introduction to the Eigenvalue Problem ....................... 261
  10.1 Matrix Review ............................................ 261
  10.2 Orthogonal matrices and the rotational operator ......... 264
  10.3 Eigenvalues and Eigenvectors ............................ 265
  10.4 Physical applications .................................... 270

11 Special functions ............................................... 279
  11.1 The factorial and gamma function ......................... 279
  11.2 The Beta Function ....................................... 283
  11.3 Stirling’s Formula ....................................... 285
  11.4 The Error Function ...................................... 288
  11.5 Elliptic Integrals ........................................ 292
  11.6 The Dirac delta function ................................ 294
**CONTENTS**

12 Power Series Solutions to Differential Equations 303
- 12.1 Power series substitution 303
- 12.2 Orthogonal vectors and Dirac Notation 307
- 12.3 Complete Sets of Functions 310
- 12.4 The Legendre Differential Equation 313
- 12.5 The Legendre Polynomials 316
- 12.6 The Generating Function for the Legendre Polynomials 319
- 12.7 Legendre Series 325
- 12.8 The Associated Legendre Differential Equation 329
- 12.9 Spherical Harmonics and the addition theorem 333
- 12.10 The Method of Frobenius and the Bessel equation 341
- 12.11 The orthogonality of the Bessel Functions 348
- 12.12 Fuch's theorem for 2-nd DE 357

13 A partial differential equation 359
- 13.1 PDE in physics 359
- 13.2 Laplace's equation in Cartesian coordinates 361
- 13.3 Laplace's equation in spherical coordinates 368
- 13.4 Laplace's equation in cylindrical coordinates 376
- 13.5 Poisson's Equation 384

14 Functions of Complex Variables 391
- 14.1 Complex variables and functions 391
- 14.2 Analytic Functions 393
- 14.3 Important Terminologies 395
- 14.4 Contour Integration and Cauchy's Theorem 400
- 14.5 Cauchy's Integral formula 402
- 14.6 Laurent's Theorem 410
- 14.7 Residues and the Residue Theorem 411
- 14.8 Methods of finding residues 413
- 14.9 Applications of the Residue Theorem 417
- 14.10 The Kramers-Kronig Relations 425
- 14.11 The revised Residue Theorem 426

15 Laplace transform 439
- 15.1 Inverse Laplace Transform 444
- 15.2 Applications of Laplace transforms 447

16 Fourier series and transform 455
- 16.1 Introduction to Fourier Series 455
- 16.2 The Fourier Series Expansion of a Periodic Function 458
- 16.3 Dirichlet Conditions 461
- 16.4 Fourier series with spatial and temporal arguments 463
- 16.5 The Fourier Transform 466
- 16.6 The Dirac Delta Function and the Fourier transform 471
- 16.7 Applications of the Fourier Transform 473
III  Mathematical Methods in physics III  483

17 Manifolds  485
17.1 What is a Manifold?  485
17.2 Curves and surfaces in a Manifold  487
17.3 Coordinate transformations and summation convention  490
17.4 The Riemannian geometry  494
17.5 Intrinsic and extrinsic geometry and the metric  497
17.6 Length, area, and volume  504
17.7 Local Cartesian coordinates and tangent space  512
17.8 The signature of a manifold  517

18 Vector Calculus on manifolds  521
18.1 The tangent vector  521
18.2 The basis vectors  522
18.3 The metric function and coordinate transformations  523
  18.3.1 Raising and lowering vector indices  529
18.4 The inner product and null vectors  530
18.5 The affine connections  531
18.6 Local geodesic and Cartesian coordinates  539
18.7 The gradient, the divergence, the curl on a manifold  542
18.8 Intrinsic derivative of a vector along a curve  544
18.9 Parallel transport  545
18.10 Null curves, non-null curves, and affine parameter  546
18.11 Refreshment from theoretical physics: the calculus of variation  548
18.12 The geodesic  550
18.13 Stationary property of the non-null geodesic  552

19 Tensor Calculus on manifolds  555
19.1 Tensors fields and rank of a tensor  555
19.2 Mapping tensors into tensors  560
19.3 Elementary tensor operations  561
19.4 Tensors and coordinate transformations  563
19.5 Tensor equations and the quotient theorem  564
19.6 Covariant derivatives of a tensor  565
19.7 Intrinsic derivative  567
Preface

This material is built upon the lecture I have given for Topics in Theoretical Physics I and II (PHYS 3150 and 3160) at Middle Tennessee State University for more than a decade.
Introduction
INTRODUCTION
Part I

Mathematical Methods in Physics I
Device 1

Series and Convergence

This lecture is based on the assumption that you have some basic knowledge of sequence, series, and tests for convergence of series. Therefore, I will only summarize the basics with more emphasis on examples.

1.1 Sequence and series

A sequence: a sequence is an ordered list of objects (or events). Like a set, it contains members (also called elements, or terms), and the number of ordered element (possibly infinite) is called the length of the sequence. Unlike a set, order matters, and exactly the same elements can appear multiple times at different positions in the sequence. A sequence is a discrete function.

A series: The sum of terms of a sequence is a series. More precisely, if \((x_1, x_2, x_3, \ldots)\) is a sequence, one may consider the sequence of partial sums \((S_1, S_2, S_3, \ldots, S_n, \ldots)\), with

\[
S_n = x_1 + x_2 + x_3 + \ldots + x_n = \sum_{k=0}^{n} x_k. \tag{1.1}
\]

Geometric series: Let’s consider two simple idealistic biological and physical processes that can be described by geometric progressions.

i. bacteria in a culture doubles every hour

\[
2, 4, 8, 16, 32, \ldots, \tag{1.2}
\]

ii. A bouncing ball that raises each time 2/3 of the height of the previous bounce:

\[
1, \frac{2}{3}, \frac{4}{9}, \frac{8}{27}, \frac{16}{81}, \ldots \tag{1.3}
\]

Suppose we want to determine the total number of bacteria or the total height the ball raises in a given period of time we add the terms in the above expressions. If we denote the sum of those terms by \(S_n\), we may write the above
expressions as

\[ S_n = 2 \left[ 1 + 2 + 2^2 + 2^3 + ... 2^n \right] \] (1.4)

\[ S_n = 1 + \frac{2}{3} + \left( \frac{2}{3} \right)^2 + \left( \frac{2}{3} \right)^3 + ... \left( \frac{2}{3} \right)^n \] (1.5)

Using the summation sign, \( \sum \), these can be put in the form

\[ S_n = \sum_{k=0}^{n} 2^k = 2 \sum_{k=0}^{n} r^k, \]

\[ S_n = \sum_{k=0}^{n} \left( \frac{2}{3} \right)^k = \sum_{k=0}^{n} r^k; \] (1.6)

where \( r = 2 \) for the doubling bacteria and \( r = 2/3 \) for the bouncing ball. In general as \( n \to \infty \), the series

\[ S = \lim_{n \to \infty} S_n = a \sum_{k=0}^{\infty} r^k \] (1.7)

is known as geometric series. Using

\[ \sum_{k=0}^{\infty} r^k = 1 + r + r^2 + r^3 + r^4 + ... \] (1.8)
1.1. SEQUENCE AND SERIES

and

\[ \sum_{k=0}^{\infty} r^{k+1} = r + r^2 + r^3 + r^4 + \ldots \]  
(1.9)

we can write

\[ \sum_{k=0}^{\infty} r^k - \sum_{k=0}^{\infty} r^{k+1} = (1 - r) \sum_{k=0}^{\infty} r^k \]

\[ = (1 + r + r^2 + r^3 + r^4 + \ldots) - (r + r^2 + r^3 + r^4 + \ldots) = 1. \]  
(1.10)

Then for \(|r| < 1\), we have

\[ \sum_{k=0}^{\infty} r^k = \frac{1}{1 - r} \]  
(1.11)

so that the geometric series becomes

\[ S = \lim_{n \to \infty} S_n = a \sum_{k=0}^{\infty} r^k = \frac{a}{1 - r}. \]  
(1.12)

Example 1.1 Two trains move toward one another with a speed \(v_o\). A bird (an “ideal physics bird”) flies back and forth between the two trains with a speed \(v_{bird} = \alpha v_o\), where \(\alpha > 1\). Find the total distance traveled by the bird before the trains crash if they start a distance \(D_0\) apart.

Solution: Consider the first four consecutive back and forth movement of the bird as shown in the figure below. The bird starts from train one \((T_1)\) and moves towards train two \((T_2)\). Let the time it takes to reach to \(T_2\) is \(t_1\) and during this time it has traveled a distance \(x_1\) as measured from the initial position of \(T_1\). Then we can express the distance \(x_1\)

\[ x_1 = D_0 - v_0 t_1. \]  
(1.13)

Since the bird is moving with a constant speed \(\alpha v_o\), the distance the bird traveled, \(x_1\), over the time, \(t_1\), is related to its speed by

\[ \alpha v_o = \frac{x_1}{t_1} \Rightarrow t_1 = \frac{x_1}{\alpha v_o}. \]  
(1.14)

Substituting Eq. (1.14) into Eq. (1.13), we find

\[ x_1 = D_0 - \frac{x_1}{\alpha} \Rightarrow x_1 = \frac{\alpha}{1 + \alpha} D_0. \]  
(1.15)

From Fig. 1.3, we note that

\[ D_1 = D_0 - 2(D_0 - x_1) = 2x_1 - D_0 \Rightarrow D_1 = 2 \left( \frac{\alpha}{1 + \alpha} \right) D_0 - D_0 \]

\[ \Rightarrow D_1 = \frac{\alpha - 1}{\alpha + 1} D_0. \]  
(1.16)
Following a similar procedure and referring to Fig. 1.3, the distance the bird traveled when it flies back to the first train, $x_2$, can be expressed as

$$x_2 = D_1 - v_0 t_2$$

(1.17)

so that using

$$\alpha v_o = \frac{x_2}{t_2} \Rightarrow t_2 = \frac{x_2}{\alpha v_o}$$

(1.18)

we find

$$x_2 = D_1 - v_0 t_2 = D_1 - \frac{x_2}{\alpha} \Rightarrow x_2 = \frac{\alpha}{1 + \alpha} D_1.$$  

(1.19)

Substituting Eq. (1.16) into Eq. (1.19), one finds

$$x_2 = \left( \frac{\alpha}{1 + \alpha} \right) \left( \frac{\alpha - 1}{\alpha + 1} \right) D_0.$$  

(1.20)

In a similar way for $x_3$, we can easily show that

$$x_3 = \frac{\alpha}{1 + \alpha} D_2.$$  

(1.21)

Using

$$D_2 = \left( \frac{\alpha - 1}{\alpha + 1} \right) D_1 = \left( \frac{\alpha - 1}{\alpha + 1} \right)^2 D_0$$

(1.22)
we find
\[ x_3 = \left( \frac{\alpha}{1 + \alpha} \right) \left( \frac{\alpha - 1}{\alpha + 1} \right)^2 D_0. \] (1.23)

Then the total distance the bird flies before the two trains crash can be expressed as an infinite series,
\[ X = x_1 + x_2 + x_3 + \ldots = \frac{\alpha}{1 + \alpha} D_0 + \left( \frac{\alpha}{1 + \alpha} \right) \left( \frac{\alpha - 1}{\alpha + 1} \right) D_0 \]
\[ + \left( \frac{\alpha}{1 + \alpha} \right) \left( \frac{\alpha - 1}{\alpha + 1} \right)^2 D_0 \ldots \]
\[ \Rightarrow X = \frac{\alpha}{1 + \alpha} D_0 \left[ \left( \frac{\alpha - 1}{\alpha + 1} \right)^0 + \left( \frac{\alpha - 1}{\alpha + 1} \right)^1 + \left( \frac{\alpha - 1}{\alpha + 1} \right)^2 + \ldots \right], \] (1.24)

which can be put in the form
\[ X = a \sum_{n=0}^{\infty} r^n, \] (1.25)

where
\[ a = \frac{\alpha}{1 + \alpha} D_0, \quad r = \frac{\alpha - 1}{\alpha + 1} < 1. \] (1.26)

Eq. (1.25) is a geometric series and its sum can easily be determined using the relation in Eq. (1.12)
\[ X = a \sum_{k=0}^{\infty} r^k = \frac{a}{1 - r} = \frac{\alpha}{1 + \alpha} D_0 \left( 1 - \frac{\alpha - 1}{\alpha + 1} \right) = \frac{1}{2} \alpha D_0. \] (1.27)

The result in Eq. (1.27) is the total distance that the bird has traveled before the two trains crashed.

1.2 Testing series for convergence

If an infinite series, like the geometric series we studied in the previous section, has a finite sum it is said to be convergent otherwise it is divergent. There are different ways of testing whether a given infinite series,
\[ a_0 + a_1 + a_2 + \ldots = \sum_{n=0}^{\infty} a_n, \] (1.28)

is a convergent or divergent series. Here we discuss some of the commonly used to test the convergence of an infinite series

1. **Preliminary test**: If \( a_n \) does not tend to zero as \( n \to \infty \),
\[ \lim_{n \to \infty} a_n \neq 0, \] (1.29)

the infinite series is a divergent series.
(2) Absolute convergence test: Suppose the sum of the absolute values of all the terms in the infinite series in Eq. (1.28) converges, then the infinite series is called absolutely convergent series. An infinite series that is absolutely convergent is a convergent series. In other words, if

$$|a_0| + |a_1| + |a_2| + ... = \sum_{n=0}^{\infty} |a_n|$$  \hspace{1cm} (1.30)

is a convergent series, the infinite series in Eq. (1.28) is a convergent series.

(3) The comparison test: Suppose there exists a convergent infinite series with all the terms positive,

$$\sum_{n=0}^{\infty} b_n = b_0 + b_1 + b_2 + ..., \hspace{1cm} (1.31)$$

(i.e. \(b_n > 0\)). Compare the \(n^{th}\) term in Eq. (1.31) \((b_n)\) with the corresponding \(n^{th}\) term in Eq. (1.28) \((a_n)\). If \(|a_n| \leq b_n\) for all \(n\), then the series in Eq. (1.28) is a convergent series.

(4) The integral test: In the series in Eq. (1.28), if \(0 < a_{n+1} \leq a_n\) for \(n > N\), then one can use the integral test,

$$I = \int_{a_n}^{\infty} a_n \, dn.$$  \hspace{1cm} (1.32)

The series in Eq. (1.28) is convergent if this integral is finite otherwise it is a divergent series. In Eq. (1.32), only the upper limit of the integration is included. This means the integral need to be evaluated only at the upper limit to test the convergence.

(5) The ratio test: In the ratio test, for the infinite series in Eq. (1.28), we first evaluate the ratio defined by

$$\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right|.$$  \hspace{1cm} (1.33)

The convergence of the series will then be determined by the value of the ratio, \(\rho\), according to

$$\rho = \begin{cases} < 1, & \text{the series is convergent,} \\ > 1, & \text{the series is divergent,} \\ = 1 & \text{use a different test.} \end{cases}$$  \hspace{1cm} (1.34)

**Example 1.2** Consider the infinite series given by

$$1 + \frac{1}{6} + \frac{2}{120} + \frac{6}{5040} + \frac{24}{362,880}, ..., \hspace{1cm} (1.35)$$
1.3. SERIES REPRESENTATIONS OF REAL FUNCTIONS

(a) Rewrite this series using standard summation notation.

(b) Use the ratio test to check the convergence of this series.

Solution:

(a) We first rewrite the series using some general expression in terms of \( n \) that generates all the terms in the series. In most cases we look for some kind of relation that can be expressed using factorial and/or exponential function. To this end, we note that

\[
1 + \frac{1}{6} + \frac{2}{120} + \frac{6}{5040} + \cdots = \frac{1}{0!} + \frac{1}{3!} + \frac{2 \times 1}{5!} + \frac{3 \times 2}{7!} + \frac{4 \times 3 \times 2}{9!} + \cdots
\]

\[
= \frac{0!}{1!} + \frac{1!}{3!} + \frac{2!}{5!} + \frac{3!}{7!} + \frac{4!}{9!} \cdots n! \frac{n!}{(2n + 1)!},
\]

so that the series can be expressed as

\[
1 + \frac{1}{6} + \frac{2}{120} + \frac{6}{5040} + \cdots = \sum_{n=0}^{\infty} a_n,
\]

where

\[
a_n = \frac{n!}{(2n + 1)!}.
\]

(b) We use the ratio test to determine if the series is convergent or divergent. Using the result determined in (a), Eqs. (1.36) and (1.37), the ratio, \( \rho \),

\[
\rho = \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \to \infty} \left| \frac{(n+1)!}{(2n+3)(2n+2)} \right| = \lim_{n \to \infty} \left| \frac{(n+1)(n!)}{(2n+3)(2n+2)(2n+1)!} \right|
\]

\[
= \lim_{n \to \infty} \left| \frac{n+1}{(2n+3)(2n+2)} \right| \Rightarrow \rho = \lim_{n \to \infty} \left| \frac{n}{4n^2} \right| = \lim_{n \to \infty} \left| \frac{1}{4n} \right| = 0
\]

which shows, \( \rho < 1 \), and the series is a convergent series according to Eq. (1.34).

1.3 Series representations of real functions

This section introduces to how we express a given well defined real function in a series form. We focus on three form of series representation of a function \( P(x) \): Power series, Taylor, and Maclaurin’s series. The series representation of this function could be convergent or divergent. The convergence or divergence is determined by the domain of the variable, \( x \), that the function depends on. We will see how to determine the range of the domain for which the series is convergent (i.e. the interval of convergence).
i. **Power Series**: A series in which the \( n^{th} \) term is a constant, \( b_n \), times the function \((x-a)^n\),

\[
P(x) = \sum_{n=0}^{\infty} b_n (x-a)^n = b_0 + b_1(x-a) + b_2(x-a)^2 + \ldots \quad (1.39)
\]

where \( a \) is a real constant and \( x \) is a real variable (i.e. \( x \in \mathbb{R} \)). Note that a power series can be convergent or divergent. It depends on the variable \( x \).

ii. **Taylor series**: Taylor series is a power series in which the coefficient in the \( n^{th} \) term of the series, \( b_n \), is given by the \( n^{th} \) derivative of the function \( P(x) \) evaluated at \( x = a \), and divided by \( n \) factorial,

\[
P(x) = \sum_{n=0}^{\infty} b_n (x-a)^n, \quad \text{where } b_n = \frac{1}{n!} \frac{d^n P(x)}{dx^n} \bigg|_{x=a} \quad (1.40)
\]

iii. **Maclaurin’s Series**: Maclaurin’s Series is a Taylor series in which the constant, \( a = 0 \):

\[
P(x) = \sum_{n=0}^{\infty} b_n x^n, \quad \text{where } b_n = \frac{1}{n!} \frac{d^n P(x)}{dx^n} \bigg|_{x=0} \quad (1.41)
\]

Any function, \( f(x) \), that is differentiable for all values of \( x \) in the specified domain, can be expressed in Taylor or Maclaurin’s series. That means if \( \frac{d^n f(x)}{dx^n} \) exists for all \( n \geq 0 \) and \( x \in \mathbb{R} \) domain, we can write

\[
f(x) = \sum_{n=0}^{\infty} b_n (x-a)^n, \quad \text{where } b_n = \frac{1}{n!} \frac{d^n f(x)}{dx^n} \bigg|_{x=a} \quad (1.42)
\]
1.3. SERIES REPRESENTATIONS OF REAL FUNCTIONS

Taylor Series for some common functions:

\[ \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \]
\[ = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots \text{ convergent for all } x \in \mathbb{R} \] \hspace{1cm} (1.43a)

\[ \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \]
\[ = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots \text{ convergent for all } x \in \mathbb{R} \] \hspace{1cm} (1.43b)

\[ e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n \]
\[ = 1 + x + \frac{x^2}{2!} + \ldots \text{ convergent for all } x \in \mathbb{R} \] \hspace{1cm} (1.43c)

\[ \ln(1 + x) = \sum_{n=1}^{\infty} \frac{(-1)^n+1}{n} x^n \]
\[ = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots \text{ convergent for } -1 < x \leq 1; \] \hspace{1cm} (1.43d)

\[ (1 + x)^p = \sum_{n=1}^{\infty} \binom{p}{n} x^n \]
\[ = 1 + px + \frac{p(p-1)}{2!} x^2 + \ldots \text{ convergent for all } |x| < 1 \] \hspace{1cm} (1.43e)

where

\[ \binom{p}{n} = \frac{p(p-1)(p-2)(p-3) \ldots (p-n+1)}{n!} \] \hspace{1cm} (1.44)

and it is called binomial coefficient.

Power series expansions of functions are unique. That is, if you use different methods for finding a power-series expansion of a function, then you must get the same result. A power series expansion for a function must be the power series expansion for that function!

**Example 1.3**

(a) Show that the Maclaurin’s Series expansion for the function

\[ f(x) = \ln(1 + x) \] \hspace{1cm} (1.45)

is expressible as

\[ \ln(1 + x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n. \] \hspace{1cm} (1.46)

(b) Find the interval of convergence for this series.
Solution:

(a) Using Eq. (1.41)

\[ P(x) = \sum_{n=0}^{\infty} b_n x^n, \text{ where } b_n = \frac{1}{n!} \left. \frac{d^n P(x)}{dx^n} \right|_{x=0} \]  

we have

\[ b_0 = \frac{1}{0!} \ln(1 + x)|_{x=0} \Rightarrow b_0 = 0, \]

\[ b_1 = \frac{1}{1!} \left. \frac{d}{dx} \ln(1 + x) \right|_{x=0} = \frac{1}{1 + x}|_{x=0} \Rightarrow b_1 = 1, \]

\[ b_2 = \frac{1}{2!} \left. \frac{d^2}{dx^2} \ln(1 + x) \right|_{x=0} = \frac{1}{2} \left. \frac{d}{dx} \left( \frac{1}{1 + x} \right) \right|_{x=0} \]

\[ = \frac{1}{2!} \frac{-1}{(1 + x)^2} \Rightarrow b_2 = \frac{(-1)^1}{2}, \]

\[ b_3 = \frac{1}{3!} \left. \frac{d^3}{dx^3} \ln(1 + x) \right|_{x=0} = \frac{1}{3!} \left. \frac{d}{dx} \left( \frac{-1}{(1 + x)^2} \right) \right|_{x=0} \]

\[ = \frac{1}{3 \times 2!} \frac{(-1)^2 (1 \times 2)}{(1 + x)^3} \Rightarrow b_3 = \frac{(-1)^2}{3}, \]

and

\[ b_4 = \frac{1}{4!} \left. \frac{d^4}{dx^4} \ln(1 + x) \right|_{x=0} = \frac{1}{4!} \left. \frac{d}{dx} \left( \frac{-1}{(1 + x)^2} \right) \right|_{x=0} \]

\[ = \frac{1}{4 \times 3!} \frac{(-1)^3 (1 \times 2 \times 3)}{(1 + x)^4} \Rightarrow b_4 = \frac{1}{4 \times 3!} \frac{(-1)^3 3!}{(1 + x)^4} \Rightarrow b_4 = \frac{(-1)^2}{4}, \]

so that one can easily see that

\[ b_n = \begin{cases} \frac{(-1)^{n+1}}{n}, & \text{for } n = 1, 2, 3... \\ 0, & n = 0 \end{cases} \]

(1.53)

There follows that

\[ \ln(1 + x) = \sum_{n=0}^{\infty} b_n x^n \]

\[ = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n. \]

(1.54)

This is the Maclaurin’s Series.
1.3. SERIES REPRESENTATIONS OF REAL FUNCTIONS

(a) To find the interval of convergence we use the ratio test. The series converges when

\[ \rho = \lim_{n \to \infty} \left| \frac{a_{n+1}(x)}{a_n(x)} \right| < 1. \] (1.55)

Using

\[ a_{n+1}(x) = \frac{(-1)^{n+2}}{n+1} x^{n+1} \] (1.56)

and

\[ a_n(x) = \frac{(-1)^{n+1}}{n} x^n, \] (1.57)

we find

\[ \rho = \lim_{n \to \infty} \left| \frac{(-1)^{n+2}}{n+1} \frac{n}{(-1)^{n+1} x^n} \right| = \lim_{n \to \infty} \left| -\frac{nx}{n+1} \right| = |x| < 1 \]

\[ \Rightarrow -1 < x < 1 \] (1.58)

We note that at \( x = \pm 1 \), the series becomes

\[ \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = \left\{ \begin{array}{ll}
\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}, & \text{for } x = 1 \\
-\sum_{n=1}^{\infty} \frac{1}{n}, & \text{for } x = -1 \end{array} \right. \] (1.59)

and the ratio test can not tell us whether the series is a convergent or divergent series. So one must use a different test. A simple integral test

\[ \int_{1}^{\infty} \frac{dn}{n} = \ln(n)|_{1}^{\infty} = \infty \] (1.60)

shows that for \( x = -1 \), the series

\[ \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = -\sum_{n=1}^{\infty} \frac{1}{n} \] (1.61)

is a divergent series. However, for \( x = 1 \) the series becomes an alternating series

\[ \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \] (1.62)

which is a convergent series.

Exercise: Prove that an alternating series of the form

\[ \sum_{n=1}^{\infty} (-1)^{n+1} a_n \] (1.63)

with \( a_n > 0 \) is a convergent series if \( a_n \) is monotonically decreasing.
Therefore, the interval of convergence for the Maclaurin’s Series is $-1 < x \leq 1$.

**Example 1.4** In Special Relativity the total energy of an object of (rest) mass $m_0$ moving with a speed $v$ is given by $E = \gamma m_0 c^2$, where $\gamma = 1/\sqrt{1 - \beta^2}$, $\beta = v/c$, and $c = 3.0 \times 10^8 \text{ m/s}$ is the speed of light in a vacuum. This total energy is the sum of the kinetic energy $T$ and the rest-mass energy $E_o = m_0 c^2$ (this is the object’s energy when the object is not moving, so $v = 0$ and $\gamma = 1$):

$$E = T + E_o = T + m_0 c^2 = \gamma m_0 c^2. \quad (1.64)$$

Show that the relativistic kinetic energy as deduced from these equations reduces to the expected (classical) form in the non-relativistic limit—that is, for $v << c$. That means we want to show that the kinetic energy of the object is given by

$$T \simeq \frac{1}{2} m_0 v^2 \quad (1.65)$$

for $v << c$ which is what you know from Intro. Physics.

**Solution:** The relativistic kinetic energy, in terms of the kinetic energy and the total energy can be expressed as

$$T = \gamma m_0 c^2 - m_0 c^2 = m_0 c^2 \left[ (1 - \beta^2)^{-1/2} - 1 \right] \quad (1.66)$$

Applying the series expansion

$$(1 + x)^p = \sum_{n=1}^{\infty} \binom{p}{n} x^n = 1 + px + \frac{p(p-1)}{2!} x^2 ... \quad (1.67)$$

for $x = -\beta^2$ and $p = -1/2$, we may write

$$(1 - \beta^2)^{-1/2} = 1 + \frac{1}{2} \beta^2 + \frac{-\frac{1}{2}(-\frac{1}{2} - 1)}{2!} \beta^4 + ... \quad (1.68)$$

In the classical limit, $v << c$ $\Rightarrow$ $\beta = v/c << 1$, we can drop all the terms with $\beta^n$ for $n \geq 4$,

$$(1 - \beta^2)^{1/2} \simeq 1 + \frac{1}{2} \beta^2 = 1 + \frac{1}{2} \frac{v^2}{c^2}. \quad (1.69)$$

Therefore, the relativistic kinetic energy in the classical limit becomes

$$T \simeq m_0 c^2 \left[ 1 + \frac{1}{2} \frac{v^2}{c^2} - 1 \right] \Rightarrow T \simeq \frac{1}{2} m_0 v^2. \quad (1.70)$$

**Example 1.5** Two identical small balls of charge $Q$ and mass $m$ are hung from two silk threads of length $L$. The threads are attached at a common point on the ceiling. The balls hang in equilibrium separated by a distance
1.3. SERIES REPRESENTATIONS OF REAL FUNCTIONS

x. Show that, for small charges \( Q \) and large lengths \( L \), the equilibrium position is given by

\[
x = \left( \frac{Q^2 L}{2\pi \varepsilon_0 mg} \right)^{1/3}.
\] (1.71)

**Solution:** If the balls are at equilibrium the net force acting on each ball must be zero. There are three forces acting on each of the balls. These are the repulsive electrostatic force \( F_e \), the tension force \( T \), and earth’s gravitational force \( W \) [Fig. 1.4]. The magnitude of the electrical \( F_e \)

![Figure 1.4: Two charged balls suspended from a massless string.](image)

and the gravitational \( W \) forces are given by

\[
F_e = \frac{1}{4\pi \varepsilon_0} \frac{Q^2}{x^2} \quad \text{and} \quad W = mg,
\]

respectively. Since the balls are at equilibrium, the net force on each ball must add up to zero. Considering the ball on the right side in Fig. 1.4, for the net for in the y-direction, we have

\[
F_{1y} = T_y - W = T \sin(\theta) - mg = 0 \Rightarrow T \sin(\theta) = mg \quad (1.72)
\]

and in the x direction

\[
F_{1x} = F_e - T_y = \frac{1}{4\pi \varepsilon_0} \frac{Q^2}{x^2} - T \cos \theta \Rightarrow T \cos \theta = \frac{1}{4\pi \varepsilon_0} \frac{Q^2}{x^2} \quad (1.73)
\]

Dividing Eq. (1.72) by (1.73), we find

\[
\tan \theta = \frac{\sin(\theta)}{\cos \theta} = \frac{4\pi \varepsilon_0 mg \pi x^2}{Q^2}.
\] (1.74)
In terms of the length of the threads, $L$, and the separation distance, $x$, we may write

$$\tan \theta = \frac{\sqrt{L^2 - \frac{x^2}{4}}} {x/2} = \frac{2L}{x} \left( 1 - \frac{x^2}{4L^2} \right)^{1/2}. \quad (1.75)$$

Applying the series expansion

$$(1 + x)^p = \sum_{n=0}^{\infty} \binom{p}{n} x^n = 1 + px + \frac{p(p-1)x^2}{2!} + \cdots \quad (1.76)$$

we can express Eq. (1.75) as

$$\frac{2L}{x} \left( 1 - \frac{x^2}{4L^2} \right)^{1/2} \approx \frac{2L}{x} \left[ 1 - \frac{1}{2} \left( \frac{x^2}{4L^2} \right) + \frac{1}{2} \frac{1}{2} \left( \frac{x^2}{4L^2} \right)^2 + \cdots \right]. \quad (1.77)$$

For large $L \gg x$ we can drop all the terms in the series that involve $L$ and approximate Eq. (1.77) as

$$\tan \theta = \frac{2L}{x} \left( 1 - \frac{x^2}{4L^2} \right)^{1/2} \approx \frac{2L}{x}. \quad (1.78)$$

Substituting Eq. (1.78) into Eq. (1.74) and solving for $x$, we find

$$\tan \theta = \frac{4\pi \epsilon_0 mg x^2}{Q^2} \Rightarrow \frac{2L}{x} \approx \frac{4\pi \epsilon_0 mg x^2}{Q^2} \Rightarrow x \approx \left( \frac{Q^2 L}{2\pi \epsilon_0 mg} \right)^{1/3}. \quad (1.79)$$

Another simple way: For the figure shown below we may write

$$F_{1y} = T_0 - W = 0 \Rightarrow T \cos \theta = mg \quad (1.80)$$

and

$$F_{1x} = F_e - T_y = 0 \Rightarrow T \sin \theta = \frac{1}{4\pi \epsilon_0} \frac{Q^2}{x^2} \quad (1.81)$$

which leads to

$$\frac{\sin \theta \cos \theta}{\cos \theta} = \frac{Q^2}{4\pi \epsilon_0 mg x^2}. \quad (1.82)$$

Apply the Maclaurin's series expansion

$$\sin \theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots \quad (1.83)$$

$$\cos \theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots \quad (1.84)$$

for small angle $\theta$ (which is the case for large $x << L$ which we should expect for small charge $Q$), we have

$$\sin \theta \approx \theta \quad \text{and} \quad \cos \theta \approx 1. \quad (1.85)$$
so that

\[ \sin(\theta) \simeq \frac{Q^2}{4\pi\epsilon_0 mx^2} \Rightarrow \frac{x}{2L} \simeq \frac{Q^2}{4\pi\epsilon_0 mgx^2} \Rightarrow x \simeq \left( \frac{Q^2 L}{2\pi\epsilon_0 mg} \right)^{1/3}. \]  \hspace{1cm} (1.86)
1.4 Homework Assignment 1

1. Land for sell: Imagine you are given a specific procedure to sell a big chunk of land to investors. This land is triangularly shaped plateau. The triangle can be approximated as an equilateral triangle. The size of this land, approximately, about \( \alpha \) square kilometers. The procedure for selling the land reads as follows:

(a) Connect the midpoints of the sides (of the equilateral triangle shaped) chunk of land to form a total of 4 smaller triangularly shaped chunks of land with equal sides (i.e. 4 equilateral triangles).

(b) Sell the middle triangularly shaped chunk of land.

(c) For each of the other 3 remaining unsold chunk of lands draw lines connecting the midpoints. You will form 4 small chunk of land from each. Again you sell each middle triangular chuck of land for another investor.

Find the infinite series for the total size (i.e. area) of land sold to the investors if the process described above continued indefinitely.

2. Find the limit of the sequence \( a_n \) given by

\[
a_n = \frac{n^2 + 5n^3}{2n^3 + 3\sqrt{4 + n^6}}.
\]

3. Use the ratio test to find whether the following series converge or diverge

(a) \[
\sum_{n=0}^{\infty} \frac{5^n (n!)^2}{(2n)!}.
\]

(b) \[
\sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{n!}.
\]

4. Find the interval of convergence for the series

\[
\sum_{n=0}^{\infty} \frac{(x - 1)^n}{2^n}.
\]

5. Prove that

\[
\sum_{i=1}^{n} i = \frac{n(n + 1)}{2}
\]

is true for any integer \( n > 0 \) (even or odd).
1.5 Homework Assignment 2

For the functions listed in problem 1 and 2: (a) Find the first four none zero terms of the Maclaurin's series; (b) Find the general term and write the series in summation form; (c) Check your result in (a) by computer; (d) Use a computer to plot the functions and several approximating partial sum of the series.

1. \[ f(x) = \frac{x}{\sqrt{1-x^2}} \]

2. \[ g(x) = \frac{1-x}{1+x} \]

*Hint: Express the function as sum or product of two functions.*

3. The little boy at the playground: you wanted to introduce a little boy to a swing set at the playground. One sunny summer day you took him to the playground. You put him on the swing and secure his sit belt. Then you pushed his sit in a horizontal direction with a force \( F \) so that the rope of the swing is displaced by an angle \( \theta \) and his sit is at a distance \( x \) from the vertical (see the figure below). At this instant the little boy is at equilibrium position. Let the weight of the little boy be \( W \) and the length of the rope is \( l \).

(a) Find \( F/W \) as a series of power of \( \theta \).

(b) In the playground it is not easy to estimate or measure the angle \( \theta \) but we can estimate \( x \) and \( l \). Find \( F/W \) as a series of power of \( x/l \).

4. Evaluate the following indeterminate functions using 'L'Hospital's rule

(a) \[ f(x) = \lim_{x \to \infty} \frac{\ln(x-a)}{\sqrt{x-a}}, \]

where \( a > 0 \).

(b) \[ g(x) = \lim_{x \to 1} (x-1)\ln[4(x-1)]. \]

5. Use the series you know to show that

\[ \frac{\pi^2}{3!} - \frac{\pi^4}{5!} + \frac{\pi^6}{7!} \ldots = 1. \]
Chapter 2

Complex Numbers, functions, and series

2.1 Complex numbers

If you punch the real number, $x = -1$, into your calculator and try to find its square roots, your calculator displays an error. That is because in reality the square root for a negative number does not exist. For Mathematicians and Physicists the square roots of a negative number do exist. It is just not in a set of real numbers but rather in more universal set of numbers known as complex numbers. A complex number, $z$, is defined by

$$z = a + ib,$$  \hspace{1cm} (2.1)

with $a$ and $b$ are real numbers and $i$ is know as the imaginary, which is defined as,

$$i^2 = -1 \Rightarrow \pm i = (-1)^{1/2}. \hspace{1cm} (2.2)$$

For any complex number, $z = a + ib$, $a$ is know as the real part and $b$ is known as the imaginary part,

$$\text{Re} \ z = a, \text{Im} \ z = b. \hspace{1cm} (2.3)$$

Rectangular and polar representation: For any complex number, $z = a + ib$, we can use a rectangular representation using the Cartesian coordinates $(x, y)$ or polar representation using the polar coordinates $(r, \theta)$. In rectangular representation the real part $a$ is the $x$ coordinate and the imaginary part $b$ is the $y$ coordinate,

$$a = x, b = y. \hspace{1cm} (2.4)$$

In polar coordinate

$$a = r \cos \theta, b = r \sin (\theta), \hspace{1cm} (2.5)$$
where $r$ is called the modulus (or magnitude) and $\theta$ is called the phase of the complex number $z$. Thus one can express the complex number $z = a + ib$
\[ z = x + iy = r \cos \theta + ir \sin (\theta). \quad (2.6) \]
The Cartesian and Polar coordinates, as shown in Fig. 2.1, are related by the equations
\[ x = r \cos \theta, y = r \sin (\theta), \quad (2.7) \]

\[ z = x + iy = r \cos \theta + ir \sin \theta \]

Figure 2.1: Rectangular and Polar representation of a complex number $z = x + iy$.

**Euler’s equation and exponential representation**: The Euler’s Equation, that relates an exponential function and trigonometric functions, is given by
\[ \exp (i\theta) = \cos (\theta) + i \sin (\theta). \quad (2.8) \]
We will derive Euler’s equation in section 2.2 when we introduce to series of complex function. Using Euler’s equation one can also represent the complex number in Eq.(2.6), in an exponential form as
\[ z = x + iy = r (\cos (\theta) + i \sin (\theta)) = r \exp (i\theta). \quad (2.9) \]
The magnitude, $r$, (which is also represented as $|z|$) and the phase angle $\theta$ for the complex number $z$ are related to the Cartesian coordinates $(x, y)$ by
\[ |z| = \sqrt{x^2 + y^2} = \sqrt{r^2 \cos^2 (\theta) + r^2 \sin^2 (\theta)} = r, \quad \theta = \tan^{-1}\left(\frac{y}{x}\right). \quad (2.10) \]
Note that the phase angle is always measured from the positive x-axis in a counterclockwise direction.
Complex conjugate of a complex Number: The complex conjugate of a complex number, \( z = x + iy \), is denoted by \( z^* \) and defined as
\[
 z^* = x - iy
\] (2.11)
The magnitude of a complex number, \( z \), is related to its complex conjugate by
\[
 |z| = \sqrt{zz^*} = \sqrt{(x + iy)(x - iy)} = \sqrt{x^2 + y^2}.
\] (2.12)

Example 2.1 Find the real and imaginary part of the complex number
\[
 z = \frac{3 + 4i}{3 - 4i}
\] (2.13)
Solution:
To find the real and imaginary parts first we need to write it in the form,
\[
 z = x + iy. \text{ This can be done by making the denominator real. In order to make it real we multiply it (and also the numerator) by complex conjugate,}
\]
\[
 z = \frac{3 + 4i}{3 - 4i} = \frac{(3 + 4i)(3 - 4i)}{(3 - 4i)(3 + 4i)} = \frac{9 + 12i + 12i - 16}{9 + 16} = \frac{-7 + 24i}{25}
\]
\[
 = \frac{-7}{25} + \frac{24}{25}i \Rightarrow \text{Re} z = -7/25, \text{Im} z = 24/25
\] (2.14)

Example 2.2 Find the real and imaginary part of the complex conjugate of the complex quantity
\[
 z = \frac{\sqrt{3} + 2i \sin (\theta)}{4 + 2e^{4i\theta}},
\] (2.15)
where \( \theta \) is real.

Solution:
The complex conjugate, \( z^* \), is given by replacing \( i \) by \(-i\)
\[
 z^* = \frac{\sqrt{3} - 2i \sin (\theta)}{4 + 2e^{-4i\theta}}
\] (2.16)
Applying Euler’s formula, we can write
\[
 2e^{-4i\theta} = 2 \cos (4\theta) - 2i \sin (4\theta)
\] (2.17)
so that
\[
 z^* = \frac{\sqrt{3} - 2i \sin (\theta)}{4 + 2 \cos (4\theta) - 2i \sin (4\theta)}
\] (2.18)
To find the real and imaginary part of \( z^* \) we need to make the denominator in the above expression real. We can do this by multiplying it with its
complex conjugate. The complex conjugate of \( 2 \cos (4\theta) - 2i \sin (4\theta) \) is \( 2 \cos (4\theta) + 2i \sin (4\theta) \). Thus

\[
\begin{align*}
z^* &= \frac{(\sqrt{3} - 2i \sin (\theta))(2 \cos (4\theta) + 2i \sin (4\theta))}{(2 \cos (4\theta) - 2i \sin (4\theta))(2 \cos (4\theta) + 2i \sin (4\theta))} \\
&= \frac{(\sqrt{3} - 2i \sin (\theta))(4 + 2 \cos 4\theta + 2i \sin 4\theta)}{(4 + 2 \cos 4\theta)^2 + 4 \sin^2 4\theta} \\
&= \sqrt{3} (1 + 2 \cos (4\theta)) + 4 \sin (4\theta) \sin (\theta) + 2i \sqrt{3} \sin (4\theta) - 2i \sin (\theta) (4 + 2 \cos (4\theta)) \\
&= \frac{\sqrt{3} (4 + 2 \cos (4\theta)) + 4 \sin (4\theta) \sin (\theta)}{(4 + 2 \cos (4\theta))^2 + 4 \sin^2 (4\theta)} \\
&\quad + i \frac{2 \sqrt{3} \sin (4\theta) - 2 \sin (\theta) (4 + 2 \cos (4\theta))}{(4 + 2 \cos (4\theta))^2 + 4 \sin^2 (4\theta)}.
\end{align*}
\] (2.19)

Therefore, the real and imaginary parts of \( z^* \) are

\[
\begin{align*}
\text{Re} (z^*) &= \frac{\sqrt{3} (4 + 2 \cos (4\theta)) + 4 \sin (4\theta) \sin (\theta)}{(4 + 2 \cos (4\theta))^2 + 4 \sin^2 (4\theta)}, \\
\text{Im} (z^*) &= \frac{2 \sqrt{3} \sin (4\theta) - 2 \sin (\theta) (4 + 2 \cos (4\theta))}{(4 + 2 \cos (4\theta))^2 + 4 \sin^2 (4\theta)}.
\end{align*}
\] (2.20)

Example 2.3 Consider the complex function

\[
Z(x) = 2f(x) - \frac{3}{2}ig(x)
\] (2.24)

where \( f(x) \) and \( g(x) \) are possibly complex functions of the real variable \( x \).

(a) Find an expression for the complex conjugate of the function \( Z(x) \).

(b) Evaluate \( Z^*(x) \) for \( x = 1 \) if

\[
f(x) = 3ix^2 \text{ and } g(x) = 2x + 4\sqrt{2}i
\] (2.25)

Solution:

(a) since \( f(x) \) and \( g(x) \) are possibly complex functions, the complex conjugate of \( Z(x) \) should be written as

\[
Z^*(x) = 2f^*(x) - \frac{3}{2}ig^*(x)
\] (2.26)

where \( f^*(x) \) and \( g^*(x) \) are the complex conjugates of \( f(x) \) and \( g(x) \), respectively.
(b) We are given

\[ f(x) = 3ix^2, \quad g(x) = 2x + 4\sqrt{2}i \]  

from which we find

\[ f^*(x) = -3ix^2, \quad g^*(x) = 2x - 4\sqrt{2}i \]  

Then

\[ Z^*(x) = 2f^*(x) + \frac{3}{2}ig^*(x) \]

becomes

\[ Z^*(x) = 2(-3ix^2) + \frac{3}{2}i(2x - 4\sqrt{2}i) \Rightarrow Z^*(x) = -6ix^2 + 3xi + 6\sqrt{2} \]

Now substituting \( x = 1 \), we find

\[ Z^*(x) = 6\sqrt{2} - 3i. \]

**Example 2.4** Find the values of the real quantities \( x \) and \( y \) if

\[ (5 + i)x - 5yi = 4 + 3i \]

**Solution:** To determine the values of the real quantities \( x \) and \( y \), we first need to write the left hand side in the form \( z = a + ib \)

\[ (5 + i)x - 5yi = 5x + i(x - 5y) = 4 + 3i \]

Two complex number \( z_1 \) and \( z_2 \) are equal if and only if their corresponding \( \text{Re} \) and \( \text{Im} \) parts are equal. Thus

\[ 5x = 4, \quad x - 5y = 3 \Rightarrow x = 4/5, \quad y = -11/25. \]

**Example 2.5** The Intensity of a Traveling Electromagnetic Waves: An electromagnetic wave is described by two mutually perpendicular oscillating electric (\( \vec{E} \)) and magnetic field (\( \vec{B} \)) vectors. An electromagnetic wave traveling in the positive \( z \)-direction the electric field can be expressed as

\[ \vec{E}(z,t) = E_0 \exp [i(\mathbf{k}z - \omega t)] \hat{x}, \]

where the amplitudes of the electric \( (E_0) \) and the angular frequency \( (\omega) \) are real. The wave vector \( (\mathbf{k} = k\hat{z}) \) given by

\[ \frac{k}{\lambda} = \frac{2\pi}{\lambda_0} = \frac{2\pi n}{\lambda_0} = \frac{2\pi n}{c} \hat{\mathbf{z}} = n \frac{\omega}{c} \hat{\mathbf{z}}, \]

in which \( \lambda_0 \) is the wave length, \( c \) is the speed of light in vacuum and \( \lambda = c/n \) is the wavelength in the medium, with refractive index \( n \), that the wave is traveling in. The refractive index of the medium is possibly
complex for some media such as metals and should be expressed, generally, as

\[ n = n_R + i n_I, \]  

(2.36)

where both \( n_R \) and \( n_I \) are real. Note that Eq. (2.36) shows the wave \( \vec{k} \) in Eq. (2.35) is also possibly complex and should be written as

\[ \vec{k} = \frac{n}{c} \hat{z} = \frac{\omega}{c} (n_R + i n_I) \hat{z}. \]  

(2.37)

The magnetic field vector (\( \vec{B} \)) usually determined from the electric field vector (\( \vec{E} \)) using the relation

\[ \vec{B}(z, t) = \vec{k} \times \vec{E}^*(z, t). \]  

(2.38)

The “Poynting vector” is an important physical quantity in describing electromagnetic waves. It is proportional to the cross product of the electric and magnetic field vectors,

\[ \vec{S}(z, t) = \frac{1}{\mu_0} \vec{E}(z, t) \times \vec{B}(z, t). \]  

(2.39)

The direction of Poynting vector shows the direction of propagation of the traveling wave and the magnitude gives the intensity of the wave as a function of position and time in the medium that the wave is propagating in. For a electromagnetic wave traveling in a medium where the refractive index is complex [Eq. (2.36)],

(a) find the complex conjugate of the electric field vector

(b) Find the magnetic field vector

(c) Shown that the magnitude of the Poynting vector is given by

\[ |\vec{S}(z, t)| = \sqrt{\vec{S}(z, t) \cdot \vec{S}^*(z, t)} = \frac{\sqrt{n_R^2 + n_I^2}}{\mu_0 c} E_0^2 \exp \left( -\frac{2\omega}{c} n_I z \right). \]  

(2.40)

**Solution:**

(a) Using

\[ \vec{E}(z, t) = E_0 \exp \left[ i (kz - \omega t) \right] \hat{x}, \]  

(2.41)

we may write the complex conjugate as

\[ \vec{E}^*(z, t) = E_0 \exp \left[ -i (k^* z - \omega t) \right] \hat{x}, \]  

(2.42)

Taking the complex conjugate of Eq. (2.37), we have

\[ k^* = \frac{\omega}{c} (n_R - i n_I). \]  

(2.43)
so that
\[ \vec{E}^*(z, t) = E_0 \exp \left[ -i \left( \frac{\omega}{c} (n_R - in_I) z - \omega t \right) \right] \hat{x} \]
\[ = E_0 \exp \left( -\frac{\omega}{c} n_I z \right) \exp \left[ -i \left( \frac{\omega}{c} n_R z - \omega t \right) \right] \hat{x} \] (2.44)

(b) In view of the result in part (a) and Eq. (2.37), the magnetic field becomes
\[ \vec{B}(z, t) = k \times \vec{E}^*(z, t) = \frac{\omega}{c} (n_R + in_I) \hat{z} \]
\[ \times E_0 \exp \left( -\frac{\omega}{c} n_I z \right) \exp \left[ -i \left( \frac{\omega}{c} n_R z - \omega t \right) \right] \hat{x} \]
\[ \Rightarrow \vec{B}(z, t) = \frac{\omega}{c} (n_R + in_I) E_0 \exp \left( -\frac{\omega}{c} n_I z \right) \exp \left[ -i \left( \frac{\omega}{c} n_R z - \omega t \right) \right] \hat{z} \times \hat{x} \]
\[ \Rightarrow \vec{B}(z, t) = \frac{\omega}{c} (n_R + in_I) E_0 \exp \left( -\frac{\omega}{c} n_I z \right) \exp \left[ -i \left( \frac{\omega}{c} n_R z - \omega t \right) \right] \hat{y} \] (2.45)

(c) Using the electric field
\[ \vec{E}(z, t) = E_0 \exp \left[ i \left( k z - \omega t \right) \right] \hat{x} = E_0 \exp \left[ i \left( \frac{\omega}{c} (n_R + in_I) z - \omega t \right) \right] \hat{x} \]
\[ \Rightarrow \vec{E}(z, t) = E_0 \exp \left( -\frac{\omega}{c} n_I z \right) \exp \left[ i \left( \frac{\omega}{c} n_R z - \omega t \right) \right] \hat{y}, \] (2.46)
the result in part (b) for the magnetic field, the pointing vector
\[ \vec{S} = \frac{1}{\mu_0} \vec{E} \times \vec{B} \] (2.47)
can be expressed as
\[ \vec{S} = \frac{1}{\mu_0} \left\{ E_0 \exp \left( -\frac{\omega}{c} n_I z \right) \exp \left[ i \left( \frac{\omega}{c} n_R z - \omega t \right) \right] \hat{x} \right\} \]
\[ \times \left\{ \frac{\omega}{c} (n_R + in_I) E_0 \exp \left( -\frac{\omega}{c} n_I z \right) \exp \left[ -i \left( \frac{\omega}{c} n_R z - \omega t \right) \right] \hat{y} \right\} \]
\[ \Rightarrow \vec{S} = \frac{1}{\mu_0} \frac{\omega}{c} (n_R + in_I) E_0^2 \exp \left( -2\frac{\omega}{c} n_I z \right) \left( \hat{x} \times \hat{y} \right) \] (2.48)
\[ \Rightarrow \vec{S} = \frac{1}{\mu_0} \frac{\omega}{c} (n_R + in_I) E_0^2 \exp \left( -2\frac{\omega}{c} n_I z \right) \hat{z} \] (2.49)

Then the magnitude of the pointing vector (the intensity)
\[ |\vec{S}(z, t)| = \sqrt{\vec{S}(z, t) \cdot \vec{S}^*(z, t)}, \] (2.50)
becomes
\[ |\vec{S}(z, t)| = \left| \frac{1}{\mu_0} \frac{\omega}{c} (n_R + in_I) E_0^2 \exp \left( -2\frac{\omega}{c} n_I z \right) \hat{z} \cdot \frac{1}{\mu_0} \frac{\omega}{c} (n_R - in_I) E_0^2 \right| \]
\[ \times \exp \left( -2\frac{\omega}{c} n_I z \right) \right|^{1/2} \]
\[ \Rightarrow |\vec{S}(z, t)| = \frac{\omega \sqrt{n_R^2 + n_I^2}}{\mu_0 c} E_0^2 \exp \left( -2\frac{\omega}{c} n_I z \right). \] (2.51)
This result shows that the intensity decays exponentially as the wave penetrates inside the medium.

2.2 Complex Infinite Series and the Circle of Convergence

Complex Infinite Series: An infinite series

\[ \sum_{n=0}^{\infty} z_n = z_0 + z_1 + z_2 + z_3 + \ldots + z_n + \ldots \]  (2.52)

is said to be a complex infinite series if

\[ z_n = a_n + ib_n, \]  (2.53)

where \( a_n \) and \( b_n \) are real number. An infinite complex series is convergent if and only if

\[ \sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + a_3 + \ldots a_n + \ldots = A \]  (2.54)

and

\[ \sum_{n=0}^{\infty} b_n = b_0 + b_1 + b_2 + b_3 + b_n + \ldots = B, \]  (2.55)

where \( A \) and \( B \) are real constants.

Test of convergence: For a complex infinite series the convergence can be tested using absolute convergence test. This means

(a) You can use any convergence test that involves the absolute value sign, which in this case will mean the magnitude of a complex quantity. This is applied to the full complex infinite series.

(b) Use any convergence test to separately test the convergence of the real and imaginary parts of the complex series. The full series converges only if both the real and imaginary parts of the series converge separately.

Series of Complex function: In the previous chapter we have seen a real power series

\[ \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 + a_2 + a_3 + \ldots a_n + \ldots \]  (2.56)

where \( a_n \) and \( x \) are real. A power series

\[ \sum_{n=0}^{\infty} c_n (z - z_0)^n = c_0 + c_1 (z - z_0) + c_2 (z - z_0)^2 + c_3 (z - z_0)^3 + \ldots c_n (z - z_0)^n + \ldots \]  (2.57)
2.2. COMPLEX INFINITE SERIES AND THE CIRCLE OF CONVERGENCE

is a series of complex function when \( z \) is a complex variable and \( z_0 \) or \( c_n \) are complex.

**Radius of convergence:** Consider the complex power series in Eq. (2.57). This series converges, applying the ratio test, for

\[
\rho = \lim_{n \to \infty} \left| \frac{c_{n+1}}{c_n} \frac{(z-z_0)^{n+1}}{(z-z_0)^n} \right| < 1 \Rightarrow \lim_{n \to \infty} \left| \frac{c_{n+1}}{c_n} \right| |z-z_0| < 1
\]  

(2.58)

For example, for

\[
\frac{c_{n+1}}{c_n} = \frac{n}{n+1}
\]  

(2.59)

we have

\[
\lim_{n \to \infty} \left| \frac{c_{n+1}}{c_n} \right| = 1
\]  

(2.60)

which leads to

\[
\rho = |z-z_0| < 1.
\]  

(2.61)

For \( z = x + iy \) and \( z_0 = a + ib \), we note that

\[
z - z_0 = x + iy - (a + ib) = (x-a) + i(y-b)
\]

\[
\Rightarrow |z - z_0| = \sqrt{(x-a)^2 + (y-b)^2} < 1 \Rightarrow (x-a)^2 + (y-b)^2 < 1^2.
\]  

(2.62)

We recall the equation

\[
(x-a)^2 + (y-b)^2 = R^2,
\]  

(2.63)

defines a circle with radius, \( R \), centered at \((a, b)\). Thus the series converges for all complex numbers, \( z \), inside the circle (disk) of radius \( R = 1 \) on the complex plane. This circle (disk) is known as **circle (disk) of convergence** for the infinite complex series in Eq. (2.57) and the radius of the circle (disk) is called the **radius of convergence**.

**Example 2.6** Determine the convergence of the following complex infinite series

\[
\sum_{n=0}^{\infty} z_n = 1 - \frac{i}{2} - \frac{1}{4!} + \frac{i}{6!} - \frac{1}{8!}
\]  

(2.64)

**Solution:** This series can be expressed as

\[
\sum_{n=0}^{\infty} z_n = \sum_{n=0}^{\infty} \frac{(-i)^n}{(2n)!}
\]  

(2.65)

Applying the ratio test we have

\[
\lim_{n \to \infty} \left| \frac{z_{n+1}}{z_n} \right| = \lim_{n \to \infty} \frac{(-i)^{n+1}}{(2n+1)!} \frac{(2n)!}{(-i)^n} = \lim_{n \to \infty} \frac{(-i)^{n+1}}{(2n+1)!} \frac{2n!}{(2n+1)!} = \lim_{n \to \infty} \left| \frac{Z_{n+1}}{Z_n} \right| = 0 < 1
\]  

(2.66)

which means the series is convergent.
Example 2.7 Find and sketch the circle of convergence in the complex plane for the following complex infinite series:

\[
\sum_{n=0}^{\infty} c_n z^n = \sum_{n=0}^{\infty} \frac{(z - 2 + i)^n}{2^n}
\]  

\hspace{0.5cm} (2.67)

**Solution:** A complex series

\[
\sum_{n=0}^{\infty} a_n (z - z_0)^n
\]  

is convergent when

\[
\rho = \lim_{n \to \infty} \left| \frac{a_{n+1} (z - z_0)^{n+1}}{a_n (z - z_0)^n} \right| < 1. 
\]  

(2.69)

Then using

\[
a_n (z - z_0)^n = \frac{(z - 2 + i)^n}{2^n}
\]

\[
a_{n+1} (z - z_0)^{n+1} = \frac{(z - 2 + i)^{n+1}}{2^{n+1}}
\]  

(2.70)

we can write

\[
\rho = \lim_{n \to \infty} \left| \frac{(z - 2 + i)^{n+1}}{2^{n+1}} \right| < 1
\]

\[
\Rightarrow \rho = \lim_{n \to \infty} \left| \frac{(z - 2 + i)^{n+1}}{2^{n+1}} \right| \frac{2^n}{(z - 2 + i)^n} < 1 \Rightarrow \left| \frac{(z - 2 + i)}{2} \right| < 1
\]

\[
\Rightarrow |z - 2 + i| < 2 \Rightarrow |z - (2 - i)| < 2.
\]  

(2.71)

This means the series is convergent inside a disk of radius \( R = 2 \) centered at \( z_0 = 2 - i \).

Example 2.8 Evaluate \( f(2 - i) \) and find the real and imaginary parts if

\[
f(z) = \frac{z - 1}{z}
\]  

(2.72)

**Solution:** We can write \( f(2 - i) \) as

\[
f(2 - i) = \frac{2 - i - 1}{2 - i} = \frac{1 - i}{2 - i}
\]  

(2.73)

which can be rewritten as

\[
f(2 - i) = \frac{(1 - i)(2 + i)}{(2 - i)(2 + i)} = \frac{3 - i}{5}
\]  

(2.74)

and the real and imaginary part

\[
\text{Re}(f(2 - i)) = 3/5, \text{Im}(f(2 - i)) = -1/5
\]  

(2.75)
Example 2.9 Derive

(b) Euler’s Formula

\[ e^{i\theta} = \cos(\theta) + i\sin(\theta) \]  
(2.76)

using series expansion of complex variables.

(b) the double-angle formulae for \( \sin(2\theta) \) and \( \cos(2\theta) \)

\[
\begin{align*}
\sin(2\theta) &= 2\sin(\theta) \cos(\theta), \\
\cos(2\theta) &= \cos^2(\theta) - \sin^2(\theta).
\end{align*}
\]  
(2.77)

Solution:

(a) Using the series expansion for the complex function

\[
\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + ..
\]  
(2.78)

we can write

\[
\exp(i\theta) = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} = 1 + i\theta - \frac{\theta^2}{2!} - \frac{i\theta^3}{3!} + ..
\]  
(2.79)

which can be put in the form

\[
\exp(i\theta) = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} + ... + i\theta - \frac{\theta^3}{3!} + \frac{i\theta^3}{3!} + ..
\]  
(2.80)

Applying the the series expansions for sine and cosine functions

\[
\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n-1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} ... \text{ convergent for all } x
\]  
(2.82)

\[
\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} ... \text{ convergent for all } x
\]  
(2.83)

we may write

\[
\exp(i\theta) = \cos(\theta) + i\sin(\theta)
\]  
(2.84)

(b) using the result above we may write

\[
[\exp(i\theta)]^2 = \exp(2i\theta) = \cos(2\theta) - i\sin(2\theta)
\]  
(2.85)

But we also know that

\[
[\exp(i\theta)]^2 = [\cos(\theta) - i\sin(\theta)]^2 \Rightarrow [\exp(i\theta)]^2 = \cos^2(\theta) - \sin^2(\theta)
\]

\[
-2i\sin(\theta) \cos(\theta)
\]  
(2.86)
Therefore, we can see that

\[
\cos(2\theta) - i\sin(2\theta) = \cos^2(\theta) - \sin^2(\theta) - 2i\sin(\theta)\cos(\theta)
\]

which leads to

\[
\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta), \quad \sin(2\theta) = 2\sin(\theta)\cos(\theta).
\]

### 2.3 Powers and Roots of complex numbers

The powers and roots of a complex number is easily determined by applying the Euler’s formula. Let’s consider a complex number

\[
z = x + iy = r \exp(i\theta),
\]

where

\[
r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1}\left(\frac{y}{x}\right).
\]

This complex number to the power \(n\) \((z^n)\) can then be written as

\[
z^n = (x + iy)^n = [r \exp(i\theta)]^n = r^n \exp(in\theta)
\]

\[
\Rightarrow z^n = (x + iy)^n = r^n (\cos(n\theta) + i \sin(n\theta)).
\]

Similarly, for the \(n^{th}\) roots to this complex number \((z^{1/n})\), one finds

\[
z^{1/n} = (x + iy)^{1/n} = r^{1/n} \exp(i\theta_k) = r^{1/n} (\cos(\theta_k) + i \sin(\theta_k))
\]

where

\[
\theta_k = \frac{\theta + 2\pi k}{n}, \quad \text{for } k = 0, 1, 2, \ldots, n - 1
\]

**Example 2.10** Solve the equation

\[
z^5 = 32
\]

and show the solutions on a complex plane.

**Solution:** Noting that

\[
z^5 = 32 = r \exp[\theta]
\]

where

\[
r = 2^5, \quad \theta = 0
\]

the solutions to

\[
z^5 = 32
\]

given by

\[
z_k = r^{1/5} \exp(i\theta_k) = r^{1/5} (\cos(\theta_k) + i \sin(\theta_k)),
\]
where
\[
\begin{align*}
r^{1/5} &= 2, \\
\theta_k &= \frac{\theta + 2\pi k}{n} = \frac{2\pi k}{n}, \text{ for } k = 0, 1, 2, 3, 4 \Rightarrow \theta_0 = 0, \theta_1 = \frac{2\pi}{5} = 72^\circ, \\
\theta_2 &= \frac{4\pi}{5} = 144^\circ, \theta_3 = \frac{6\pi}{5} = 216^\circ, \theta_4 = \frac{8\pi}{5} = 288^\circ \quad (2.99)
\end{align*}
\]

Then the solution to the equation are found to be

\[z_0 = 2 \left[ \cos \left( \frac{2\pi}{5} \right) + i \sin \left( \frac{2\pi}{5} \right) \right] = 2 + i0
\]
\[z_1 = 2 \left[ \cos \left( \frac{4\pi}{5} \right) + i \sin \left( \frac{4\pi}{5} \right) \right] = 0.62 + i1.90
\]
\[z_2 = 2 \left[ \cos \left( \frac{6\pi}{5} \right) + i \sin \left( \frac{6\pi}{5} \right) \right] = -1.62 + i1.18
\]
\[z_3 = 2 \left[ \cos \left( \frac{8\pi}{5} \right) + i \sin \left( \frac{8\pi}{5} \right) \right] = -1.62 - i1.18
\]
\[z_4 = 2 \left[ \cos \left( \frac{8\pi}{5} \right) + i \sin \left( \frac{8\pi}{5} \right) \right] = 1.62 - i1.90 \quad (2.100)
\]

These roots are shown on the complex plane in Fig. 2.2
2.4 Algebraic vs transcendental functions

The logarithm and the exponential function are examples of transcendental functions. Transcendental function is a term often used to describe the trigonometric functions, i.e., sine, cosine, tangent, cotangent, secant, cosecant, and so forth. A function that is not transcendental is said to be algebraic. Examples of algebraic functions are rational functions and the square root function. A composition of transcendental functions can give an algebraic function. For example

\[ f(x) = \cos(\arcsin(x)) = \sqrt{1 - x^2} \]  \hspace{1cm} (2.101)

Euler's Equation, trigonometric, and hyperbolic Functions: We recall the Euler's formula in terms of a real variable \( x \) is given by

\[ \exp(ix) = \cos(x) + i\sin(x) \]  \hspace{1cm} (2.102)

and taking the complex conjugate we have

\[ \exp(-ix) = \cos(x) - i\sin(x) \]  \hspace{1cm} (2.103)

so that adding the two equations we find

\[ \cos(x) = \frac{e^{ix} + e^{-ix}}{2} \]  \hspace{1cm} (2.104)

and subtracting

\[ \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}. \]  \hspace{1cm} (2.105)

If \( x \) is a complex variable \( z \), using Euler’s Equation, we can express the sine and cosine functions as

\[ \cos(z) = \frac{e^{iz} + e^{-iz}}{2} \]  \hspace{1cm} (2.106)

and

\[ \sin(z) = \frac{e^{iz} - e^{-iz}}{2i} \]  \hspace{1cm} (2.107)

If \( z \) is a pure imaginary complex number \( z = iy \), the sine and cosine functions become

\[ \cos(iy) = \frac{e^{-y} + e^{y}}{2} \Rightarrow \cos(iy) = \cosh(y) \]  \hspace{1cm} (2.108)

and

\[ \sin(iy) = \frac{e^{-y} - e^{y}}{2i} = i\frac{e^{y} - e^{-y}}{2} \Rightarrow \sin(iy) = i\sinh(y) \]  \hspace{1cm} (2.109)

where

\[ \cosh(y) = \frac{e^{y} + e^{-y}}{2}, \sinh(y) = \frac{e^{y} - e^{-y}}{2} \]  \hspace{1cm} (2.110)

are called the cosine hyperbolic and sine hyperbolic functions. It can be generalized for a complex number \( z \) as

\[ \sinh(z) = \frac{e^{-z} + e^{z}}{2}, \cosh(z) = \frac{e^{z} + e^{-z}}{2}. \]  \hspace{1cm} (2.111)
Other hyperbolic functions: named and defined in a way similar to the trigonometric functions:

\[
\tanh (z) = \frac{\sinh (z)}{\cosh (z)}, \quad \coth (z) = \frac{1}{\tanh (z)} \quad (2.112)
\]

\[
\csc h (z) = \frac{1}{\sinh (z)}, \quad \sec h (z) = \frac{1}{\cosh (z)}
\]

Logarithms: Consider a complex number \( z = re^{i\theta} \), then the logarithm of \( z \) to the base \( e \) which is denoted by \( \ln z \) can be expressed as

\[
\ln z = \ln (re^{i\theta}) = \ln r + i\theta \ln e \Rightarrow \ln z = \ln r + i\theta \quad (2.113)
\]

The principal value: the principal value of a complex number \( z \) expressed in polar coordinates as

\[
z = re^{i\theta} \quad (2.114)
\]

is the value of the complex number when \( 0 \leq \theta \leq 2\pi \). The principal value of \( \ln z \) is the value of \( \ln z = \ln (re^{i\theta}) \) for \( 0 \leq \theta \leq 2\pi \).

Example 2.11 Evaluate the following integral applying Euler’s formula,

\[
I = \int_{-\pi}^{\pi} \sin (\alpha x) \cos (\beta x) \, dx \quad (2.115)
\]

where \( \alpha \) and \( \beta \) are positive, nonzero integers.

Solution: Applying Euler’s formula we can write

\[
\sin (\alpha x) = \frac{e^{i\alpha x} - e^{-i\alpha x}}{2i}, \quad \cos (\beta x) = \frac{e^{i\beta x} + e^{-i\beta x}}{2} \quad (2.116)
\]

so that

\[
\sin (\alpha x) \cos (\beta x) = \left( \frac{e^{i\alpha x} - e^{-i\alpha x}}{2i} \right) \left( \frac{e^{i\beta x} + e^{-i\beta x}}{2} \right)
= \frac{e^{i(\alpha+\beta)x} - e^{-i(\alpha+\beta)x} + e^{i(\alpha-\beta)x} - e^{-i(\alpha-\beta)x}}{4i} \quad (2.117)
\]

Then substituting the above expression for the integrand, we have

\[
I = \frac{1}{4i} \int_{-\pi}^{\pi} \left[ e^{i(\alpha+\beta)x} - e^{-i(\alpha+\beta)x} + e^{i(\alpha-\beta)x} - e^{-i(\alpha-\beta)x} \right] \, dx \quad (2.118)
\]

so that applying Euler’s formula, once again, one can write

\[
e^{i(\alpha+\beta)x} - e^{-i(\alpha+\beta)x} = 2i \sin [(\alpha + \beta) x], \quad e^{i(\alpha-\beta)x} - e^{-i(\alpha-\beta)x} = 2i \sin [(\alpha - \beta) x]. \quad (2.119)
\]

Using Eq. (2.119), we can write the integral as

\[
I = \frac{1}{2} \int_{-\pi}^{\pi} [\sin [(\alpha + \beta) x] + \sin [(\alpha - \beta) x]] \, dx. \quad (2.120)
\]
which leads to
\[ I = -\frac{1}{2} \left[ \cos \left( \frac{(\alpha + \beta) x}{\alpha + \beta} \right) + \cos \left( \frac{(\alpha - \beta) x}{\alpha - \beta} \right) \right]_\pi^{-\pi} \]  (2.121)

Noting that cosine is an even function,
\[ \cos(x) = \cos(-x) \]
we can easily see that Eq. (2.121) reduces to
\[ I = \cos \left( \frac{(\alpha + \beta) \pi}{\alpha + \beta} \right) - \cos \left( \frac{(\alpha - \beta) \pi}{\alpha - \beta} \right) = 0 \]  (2.122)

Therefore
\[ I = \int_{-\pi}^{\pi} \sin(\alpha x) \cos(\beta x) \, dx = 0. \]  (2.123)

**Orthonormal Sets of Functions:** The set of functions \( \{f_1(x), f_2(x), \ldots f_j(x), \ldots \} \) is said to be orthonormal if
\[ \int_{x_1}^{x_2} f_i(x)f_j(x) \, dx = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \]  (2.124)

**Example 2.12** Evaluate the following integral, where \( \alpha \) and \( \beta \) are positive, nonzero integers:
\[ I = \int_{-\pi}^{\pi} \sin(\alpha x) \sin(\beta x) \, dx \]  (2.125)
and show that the sine function form an orthonormal set of functions.

**Solution:** Applying Euler’s formula, we may write
\[ \sin(\alpha x) = \frac{e^{i\alpha x} - e^{-i\alpha x}}{2i}, \sin(\beta x) = \frac{e^{i\beta x} - e^{-i\beta x}}{2i} \]  (2.126)
so that
\[
\sin(\alpha x) \sin(\beta x) = \left( \frac{e^{i\alpha x} - e^{-i\alpha x}}{2i} \right) \left( \frac{e^{i\beta x} - e^{-i\beta x}}{2i} \right) \\
= \frac{e^{i(\alpha+\beta)x} - e^{i(\alpha-\beta)x} - e^{-i(\alpha-\beta)x} + e^{-i(\alpha+\beta)x}}{-4} \\
= \frac{1}{2} \left\{ e^{i(\alpha-\beta)x} + e^{-i(\alpha-\beta)x} - e^{i(\alpha+\beta)x} - e^{-i(\alpha+\beta)x} \right\} \\
= \frac{1}{2} \{ \cos[(\alpha - \beta)x] - \cos[(\alpha + \beta)x] \} \]  (2.127)

This can be rewritten as
\[ \sin(\alpha x) \sin(\beta x) = \frac{1}{2} \left\{ \cos[(\alpha - \beta)x] - \cos[(\alpha + \beta)x] \right\} \]  (2.128)
or
\[ \sin(\alpha x) \sin(\beta x) = \frac{1}{2} \{ \cos[(\alpha - \beta)x] - \cos[(\alpha + \beta)x] \} \]  (2.129)
2.4. ALGEBRAIC VS TRANSCENDENTAL FUNCTIONS

Then substituting the above expression for the integrand, we have

\[ I = \frac{1}{2} \int_{-\pi}^{\pi} \{ \cos \[(\alpha - \beta) x\] - \cos \[(\alpha + \beta) x\] \} \, dx \]  
(2.130)

\[ \Rightarrow I = \frac{1}{2} \left\{ \frac{\sin \[(\alpha - \beta) x\]}{\alpha - \beta} - \frac{\sin \[(\alpha + \beta) x\]}{\alpha + \beta} \right\} \bigg|_{-\pi}^{\pi}. \]  
(2.131)

We recall \( \sin \) is an odd function, which means

\[ \sin (-x) = -\sin (x) \]  
(2.132)

we have

\[ \sin \left[ - (\alpha - \beta) \pi \right] = -\sin \left[ (\alpha - \beta) \pi \right], \sin \left[ - (\alpha + \beta) \pi \right] = -\sin \left[ (\alpha + \beta) \pi \right] \]

so that

\[ I = \frac{\sin \left[ (\alpha - \beta) \pi \right]}{\alpha - \beta} - \frac{\sin \left[ (\alpha + \beta) \pi \right]}{\alpha + \beta} \]  
(2.133)

Since \( \alpha \) and \( \beta \) are positive, none zero integers, we always have

\[ \sin \left[ (\alpha + \beta) \pi \right] = 0 \]  
(2.135)

which leads to

\[ I = \frac{\sin \left[ (\alpha - \beta) \pi \right]}{\alpha - \beta}. \]  
(2.136)

This leads to

\[ I = \int_{-\pi}^{\pi} \sin (\alpha x) \sin (\beta x) \, dx = \delta_{\alpha,\beta} = \begin{cases} \pi & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases} \]  
(2.137)

which means the sine function

\[ f_j(x) = \sin (jx) \]  
(2.138)

form an orthogonal set of functions in the domain \((-\pi, \pi)\).

**Example 2.13** Evaluate the complex expression

\[ \cos [i \ln 2] \]  
(2.139)

**Solution:** Evaluating this expression is much easier if we use the Euler’s formula:

\[ \cos (z) = \frac{e^{iz} + e^{-iz}}{2} \]  
(2.140)

we can express \( \cos [i \ln 2] \) as

\[ \cos [i \ln 2] = \frac{e^{i(i \ln 2)} + e^{-i(i \ln 2)}}{2} = \frac{e^{-\ln 2} + e^{\ln 2}}{2} \]  
(2.141)
and using the relation
\[ a^b = e^{b \ln a}, \tag{2.142} \]
for \( a = 2, b = -1 \) (or 1), we have
\[ e^{-\ln 2} = 2^{-1} = 0.5, e^{\ln 2} = 2^1 = 2 \tag{2.143} \]
so that
\[ \cos [i \ln 2] = \frac{3/2}{2} = \frac{3}{4} = 0.75. \tag{2.144} \]

**Example 2.14** Find the principal value of the complex logarithm
\[ z = \ln (1 - i) \tag{2.145} \]
**Solution:** To find the principal value first we need to express \( z = 1 - i \) using polar coordinates. To this end, we note that
\[ r = |1 - i| = \sqrt{1^2 + (-1)^2} = \sqrt{2}, \theta = \frac{7}{4} \pi + 2\pi n \tag{2.146} \]
where \( n = 0, 1, 2 \ldots \) But to find the principal value for \( \ln z \) we use \( \theta = \frac{7}{4} \pi \).
Thus
\[ z = 1 - i = \sqrt{2} e^{i \frac{7}{4} \pi} \tag{2.147} \]
which leads to
\[ \ln z = \ln 2^{1/2} + \frac{7\pi i}{4} \Rightarrow \ln z = \frac{1}{2} \ln 2 + \frac{7\pi i}{4} \tag{2.148} \]

**Example 2.15** Find the principal value of the complex quantity
\[ z = (1 - i)^i \tag{2.149} \]
**Solution:** To find the principal value we first put the complex number in the form \( re^{i\theta} \). Recalling that
\[ 1 - i = \sqrt{2} \exp \left[ i \left( \frac{7}{4} \pi + 2\pi n \right) \right] \tag{2.150} \]
we have
\[ (1 - i)^i = \left( \sqrt{2} \right)^i \exp \left[ i \left( \frac{7}{4} \pi + 2\pi n \right) \right] \tag{2.151} \]
\[ (1 - i)^i = \left( 2^{1/2} \right)^i \exp \left[ i \times i \left( \frac{7}{4} \pi + 2\pi n \right) \right] \]
\[ (1 - i)^i = 2^{\frac{i}{2}} \exp \left[ - \left( \frac{7}{4} \pi + 2\pi n \right) \right] \]
The principal value is when \( n = 0 \). That means
\[ (1 - i)^i = 2^{\frac{i}{2}} \exp \left[ - \frac{7}{4} \pi \right] \tag{2.152} \]
Example 2.16 Find the value for the following complex expression

\[ z = \sin^{-1} \left[ \left( \frac{\sqrt{3} + i}{\sqrt{3} - i} \right)^{12} \right] \]  

(2.153)

Solution: Let the value of this expression be \( \theta \) so that

\[ \sin (\theta) = \left( \frac{\sqrt{3} + i}{\sqrt{3} - i} \right)^{12} . \]  

(2.154)

Noting that

\[ \tan^{-1} \left[ \frac{1}{\sqrt{3}} \right] = \frac{\pi}{6} \Rightarrow \sqrt{3} + i = 2e^{i\frac{\pi}{6}} \]  

(2.155)

\[ \tan^{-1} \left[ \frac{-1}{\sqrt{3}} \right] = \frac{11\pi}{6} \Rightarrow \sqrt{3} - i = 2e^{i\frac{11\pi}{6}} \]  

(2.156)

we may write

\[ \sin (\theta) = \left( \frac{2e^{i\frac{\pi}{6}}}{2e^{i\frac{11\pi}{6}}} \right)^{12} = \left[ e^{i(\frac{\pi}{6} - \frac{11\pi}{6})} \right]^{12} \]  

(2.157)

\[ \sin (\theta) = e^{i12(\frac{\pi}{6} - \frac{11\pi}{6})} = e^{-20i\pi} = e^{-10i(2\pi)} = 1 \]

\[ \Rightarrow \theta = \frac{\pi}{2} \pm 2k\pi, k = 0, 1, 2... \]  

(2.158)
2.5 Homework Assignment 3

1. For the complex number 
   \( z = \sqrt{2}e^{-i\pi} \)
   (a) Find \( x, y, r, \theta \) and write it in the forms 
   \( z = (x, y), z = x + iy, z = (r, \theta), \) and \( z = r(\cos \theta + i \sin \theta) \).
   (b) Find \( z^* \) and also write it in the forms 
   \( z = (x, y), z = x + iy, z = (r, \theta), \) and \( z = r(\cos \theta + i \sin \theta) \).
   (c) Plot both \( z \) and \( z^* \) and label it.

2. Write the complex expression 
   \( z = \frac{1}{(2 - 3i)^2} \)
   in the form \( z = a + bi \).

3. Prove that 
   \( \left( \frac{z_1}{z_2} \right)^* = \frac{z_1^*}{z_2^*} \)
   (i.e. the conjugate of the quotient of two complex numbers is the quotient of the conjugates.) Also prove that 
   \( (z_1 z_2)^* = z_1^* z_2^* \)
   \( (z_1 \pm z_2)^* = z_1^* \pm z_2^* \)

   Hint: it is easier to prove the statements about the quotient and the product using the polar coordinates \( re^{i\theta} \) form; for the sum or difference it is easier to use the rectangular form \( x + iy \).

4. 
   (a) Find the absolute value of a complex number 
   \( z = (1 + 2i)^3 \)
   (b) Solve for all possible values of the real number \( x \) and \( y \) in the equation 
   \( 2ix + 3 = y - i \)

5. Physical application: A particle moves in the \((x, y)\) plane so that its position 
   \((x, y)\) as a function of time is given by 
   \( z = \cos(3t) + i \sin(3t) \)
(a) Find the velocity \((v)\) and acceleration \((a)\) of the particle

\[
v = \frac{dz}{dt}, \quad a = \frac{d^2z}{dt^2}
\]

(b) Find the magnitude (absolute value) of the velocity and acceleration

(c) Describe the motion of the particle.
2.6 Homework Assignment 4

1. Test each of the following series for convergence.

(a) \[ \sum_{n=0}^{\infty} \frac{1}{(1 + i)^n} = 1 + \frac{1}{(1 + i)} + \frac{1}{(1 + i)^2} + \ldots + \frac{1}{(1 + i)^n} + \ldots \]

(b) \[ \sum_{n=0}^{\infty} \left( \frac{1 - i}{1 + i} \right)^n = 1 + \frac{1 - i}{1 + i} + \left( \frac{1 - i}{1 + i} \right)^2 + \ldots + \left( \frac{1 - i}{1 + i} \right)^n + \ldots \]

2. Find the disk of convergence for each of the following power series of complex variable

(a) \[ \sum_{n=0}^{\infty} \frac{(n!)^3 z^n}{(3n)!} = 1 + \frac{z}{3!} + \frac{(2!)^3 z^2}{6!} + \frac{(n!)^3 z^n}{(3n)!} + \ldots \]

(b) \[ \sum_{n=1}^{\infty} \frac{(z - i)^n}{n} = \frac{z - i}{1} + \frac{(z - i)^2}{2} + \frac{(z - i)^3}{3} + \ldots + \frac{(z - i)^n}{n} + \ldots \]

3. (a) Show that for any real \( y \), \( |e^{iy}| = 1 \). Hence show that \( |e^{z}| = e^{x} \) for every complex \( z \)

\[ |e^{z}| = e^{x}. \]

(b) For the complex numbers \( z_1 \) and \( z_2 \), prove that

\[ |z_1 z_2| = |z_1||z_2| \]
\[ \frac{z_1}{z_2} = \frac{|z_1|}{|z_2|} \]

*Hint: write the numbers in re^{i\theta} form*

(c) Using the relations in part (a) and (b) find the value for

\[ W = \left| \frac{1 - i}{1 + i} \right| \]

4. Find the roots for

\[ z = i^{1/5}. \]

and show your results on the complex plane.

5. Find the formula for \( \sin(3\theta) \) and \( \cos(3\theta) \).

*Hint: refer to the example for double angle formula that we did in class.*
2.7 Homework Assignment 5

Find the values in rectangular form \( x + iy \) for the complex expressions in problem 1-4

1. 
\[ z = \sin (\pi - i \ln 3) \]

2. 
\[ z = \tan i \]

3. 
\[ z = \cos (2i \ln i) \]

4. 
\[ z = \tanh^{-1} \left( i \sqrt{3} \right) \]

5. Physical Application: Consider the RLC-ac circuit shown in the figure below. The circuit has a resistor and an inductor connected in series and then a capacitor in parallel with them (\( R \) and \( L \) in series, and then \( C \) in parallel with them). From introductory physics, the impedance for a resistor (\( Z_R \)), capacitor (\( Z_C \)), and inductor (\( Z_L \)), are given by

\[ Z_R = R, \quad Z_C = 1/(i\omega C), \quad \text{and} \quad Z_L = i\omega L, \]

where \( R \) is the resistance of the resistor, \( C \) is the capacitance of the capacitor, \( L \) is the inductance of the inductor, and \( \omega \) is the angular frequency of the ac-source voltage, \( V(t) = V_{\text{max}} \sin (\omega t) \) in the circuit. In ac circuit for series connection of two impedances, \( Z_1 \) and \( Z_2 \) the equivalence impedance \( Z_{12} \) is given by

\[ Z_{12} = Z_1 + Z_2 \]
where as for parallel connection

\[
\frac{1}{Z_{12}} = \frac{1}{Z_1} + \frac{1}{Z_2}.
\]

(a) Find the equivalent impedance of the circuit

(b) Find the magnitude of the equivalent impedance of the circuit.

(c) A circuit is said to be at resonance when the total impedance of the circuit is real. Find \( \omega \) in terms of \( R, L, \) and \( C \) at resonance.
Chapter 3

Vectors, Lines, and Planes

3.1 An Overview of Vector Fundamentals

A vector: A vector is a quantity that can be described by a magnitude and direction. We shall represent a vector by a boldface letter (for example $\mathbf{A}$) or by a letter with an arrow on top of the letter (for example $\vec{A}$). Here we will use the second way of representation. The component of a vector is represented by a subscript (for example $A_x$ represent the $x$ component of the vector $\vec{A}$).

![Figure 3.1: A two-dimensional vector and its components on the x-y plane.](image)

Magnitude of a vector: the length of the arrow representing a vector $\vec{A}$ is called the length or the magnitude of $\vec{A}$ (written $|\vec{A}|$ or $A$) and is given by

\[
A = \sqrt{A_x^2 + A_y^2} \quad (2\text{-Dimensions})
\]

\[
A = \sqrt{A_x^2 + A_y^2 + A_z^2} \quad (3\text{-Dimensions})
\]  \hspace{1cm} (3.1)
A unit vector: A unit vector is a vector whose magnitude is one. For any vector \( \vec{A} \) we can find a unit vector denoted by \( \hat{A} \) (a vector along the direction of \( \vec{A} \) with unit magnitude). This unit vector is given by

\[
\hat{A} = \frac{\vec{A}}{|\vec{A}|}
\]  

(3.2)

### 3.2 Vector Multiplication

Two vectors can be multiplied in two different ways. One way of multiplying two vectors leads to a scalar. Such multiplication of vectors is referred as Scalar product (dot product). The other way of multiplying two vectors leads to another vector and is known as cross product (vector product)

**A. The Scalar Product (The Dot Product):** The scalar product of two vectors \( \vec{A} = A_x \hat{i} + A_y \hat{j} + A_z \hat{k} \) and \( \vec{B} = B_x \hat{i} + B_y \hat{j} + B_z \hat{k} \) (written as \( \vec{A} \cdot \vec{B} \)) is given by

\[
\vec{A} \cdot \vec{B} = A_x B_x + A_y B_y + A_z B_z.
\]  

(3.3)

If the magnitude of the two vectors and the angle between the two, \( \theta \), is known, the scalar product of these two vectors can also be determined using

\[
\vec{A} \cdot \vec{B} = |\vec{A}| |\vec{B}| \cos \theta.
\]  

(3.4)

A dot product can be used to determine the angle between two vectors

\[
\theta = \cos^{-1} \left( \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|} \right) = \cos^{-1} \left( \frac{A_x B_x + A_y B_y + A_z B_z}{|\vec{A}| |\vec{B}|} \right)
\]  

(3.5)

**Application:** Suppose you pull an object on a frictionless surface by applying a force and moved the object horizontally for some distance. If the force applied is \( \vec{F} \) and the displacement of the object is \( \vec{S} \), the work done \( W \) is given by

\[
W = \vec{F} \cdot \vec{S} = |\vec{F}| |\vec{S}| \cos (\theta)
\]  

(3.6)

**B. The Vector Product (The Cross Product):** The vector product of two vectors \( \vec{A} \) and \( \vec{B} \) (written as \( \vec{A} \times \vec{B} \)) gives a third vector \( \vec{C} \) perpendicular to the plane formed by the two vectors. If the angle between the two vectors is \( \theta \), the magnitude of \( \vec{C} \) is given by

\[
C = |\vec{A} \times \vec{B}| = |\vec{A}| |\vec{B}| \sin (\theta)
\]  

(3.7)

and the direction can be determined using the right-hand rule.
3.2. VECTOR MULTIPLICATION

The Right-Hand Rule: To find the direction of \( \vec{C} = \vec{A} \times \vec{B} \) put your right hand on the first vector (\( \vec{A} \)) curl your fingers towards the second vector (\( \vec{B} \)) then your thumb points in the direction of vector \( \vec{C} \).

Right-Handed Coordinate Systems: in right-hand coordinate system the right hand rule is valid. For example in Cartesian coordinate system we have

\[
i \times j = \hat{k}, \; j \times k = \hat{i}, \; k \times i = \hat{j}.
\]  

(3.8)

If we know the components of the two vectors, \( \vec{A} = A_x\hat{i} + A_y\hat{j} + A_z\hat{k} \) and \( \vec{B} = B_x\hat{i} + B_y\hat{j} + B_z\hat{k} \), then you can find the components of the vector product \( \vec{C} = \vec{A} \times \vec{B} = C_x\hat{i} + C_y\hat{j} + C_z\hat{k} \), in Cartesian coordinates, using

\[
\begin{align*}
C_x &= A_yB_z - A_zB_y \\
C_y &= A_zB_x - A_xB_z \\
C_z &= A_xB_y - A_yB_x
\end{align*}
\]  

(3.9)

Application of cross product: torque due to a force \( \vec{F} \) on a ridged object placed at a position \( \vec{r} \) with respect to an axis (through the origin) is given by \( \vec{\tau} = \vec{r} \times \vec{F} \).

Example 1 A force of magnitude 4.0 N pointing in the positive-x direction is applied to a large solid object at the position (0, 2, 0)m. Find the resulting torque exerted on the object about the origin using the

(a) the magnitude and direction form of the cross product.

(b) component form of the cross product.

Solution:
(a) The torque is given by
\[ \vec{\tau} = \vec{r} \times \vec{F} \] (3.10)
where the force, pointing along the positive-x direction is given by
\[ \vec{F} = (4, 0, 0) \text{ N} = 4N \hat{x} \] (3.11)
and the position vector
\[ \vec{r} = (0, 2, 0) \text{ m} = 2m \hat{y} \] (3.12)
We use the relation for the magnitude of the cross product of two vectors given by
\[ |\vec{\tau}| = |\vec{r}| |\vec{F}| \sin(\theta) \] (3.13)
where \( \theta \) is the angle between the position and the force vector. The magnitude of the position and force vectors are obtained using
\[ |\vec{r}| = \sqrt{r_x^2 + r_y^2 + r_z^2} = 2m \] (3.14)
and
\[ |\vec{F}| = \sqrt{F_x^2 + F_y^2 + F_z^2} = 4N. \] (3.15)
Noting that the angle between \( \vec{r} \) and \( \vec{F} \) is
\[ \theta = \frac{\pi}{2} \] (3.16)
the magnitude of the torque would be
\[ |\vec{\tau}| = |\vec{F}| |\vec{r}| \sin(\theta) = 8Nm. \] (3.17)
To find the direction use the right hand rule and you will find the direction to be along the negative-z direction. Then the torque may be expressed as
\[ \vec{\tau} = -8Nm \hat{z} \] (3.18)
(b) The torque is given by
\[ \vec{\tau} = \vec{r} \times \vec{F} \] (3.19)
where the force, pointing along the positive-x direction is given by
\[ \vec{F} = (4, 0, 0) \text{ N} = 4N \hat{x} \] (3.20)
and the position vector
\[ \vec{r} = (0, 2, 0) \text{ m} = 2m \hat{y} \] (3.21)
We apply the relation

\[ \vec{A} \times \vec{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \]  

(3.22)

which gives

\[ \vec{r} = \vec{r}_0 \times \vec{F} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ 0 & 2 & 0 \\ 4 & 0 & 0 \end{vmatrix} = \hat{x} \begin{vmatrix} 2 & 0 \\ 0 & 4 \end{vmatrix} \hat{y} \begin{vmatrix} 0 & 0 \\ 4 & 0 \end{vmatrix} + \hat{z} \begin{vmatrix} 0 & 2 \\ 4 & 0 \end{vmatrix} \]  

(3.23)

\[ \vec{r} = \vec{r}_0 \times \vec{F} = -8 N m \hat{z} \]

### 3.3 Equations of a straight line

**Symmetric equations** (in a two dimensional space):

\[ \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c} \]  

(3.24)

where \( a, b, \) and \( c \neq 0 \). When \( c = 0 \) we should write the symmetric equation of the straight line as

\[ \frac{x - x_0}{a} = \frac{y - y_0}{b}, z = z_0 \]  

(3.25)

**Parametric equations:**

\[ \vec{r} = \vec{r}_0 + \vec{A}t \quad \text{or} \quad \begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{cases} \]  

(3.26)

The symmetric equations can be obtained from the parametric equations:

\[ t = \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c} \]  

(3.27)

### 3.4 Equation of a plane

If \( \vec{N} = ai + bj + ck \) is normal (perpendicular) to a plane, then the dot product of the vector \( \vec{N} \) and the vector \( \vec{r} - \vec{r}_0 \)

\[ \vec{r} - \vec{r}_0 = (x - x_0) \hat{x} + (y - y_0) \hat{y} + (z - z_0) \hat{z} \]  

(3.28)

is zero,

\[ a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \]  

(3.29)

This defines the equation of the plane. It can be rewritten as

\[ ax + by + cz = d \]  

(3.30)

where

\[ d = ax_0 + by_0 + cz_0 \]  

(3.31)
Example 2: The position of a particle at time $t = 0$ is given by $(1, 2, 3)m$. The particle moves with a constant velocity $(0, 2, 1)m/s$. What is the equation of the particle’s trajectory?

Solution: The initial position of the particle can be expressed, using Cartesian unit vectors, as

$$\vec{r}_0 = \hat{x} + 2\hat{y} + 3\hat{z} \quad (3.32)$$

at a later time $t$, let the position of the particle is described by the vector

$$\vec{r} = x(t) \hat{x} + y(t) \hat{y} + z(t) \hat{z} \quad (3.33)$$

If the particle is traveling by a constant velocity, the vector

$$\vec{r} - \vec{r}_0 = [x(t) - 1] \hat{x} + [y(t) - 2] \hat{y} + [z(t) - 3] \hat{z} \quad (3.34)$$

must be proportional to the velocity vector

$$\vec{v} = 0\hat{x} + (2m/s) \hat{y} + (1m/s) \hat{z} \quad (3.35)$$

This means

$$\vec{r} - \vec{r}_0 = t\vec{v} \Rightarrow [x(t) - 1] \hat{x} + [y(t) - 2] \hat{y} + [z(t) - 3] \hat{z} = (2m/s) t\hat{y} + (1m/s) t\hat{z} \quad (3.36)$$

$$\Rightarrow x(t) \hat{x} + y(t) \hat{y} + z(t) \hat{z} = 1\hat{x} + 2\hat{y} + 3\hat{z} + (2m/s) t\hat{y} + (1m/s) t\hat{z} \quad (3.37)$$

or

$$\vec{r}'(t) = \vec{r}_0 + \vec{v} t \quad (3.38)$$

where

$$\vec{r}_0 = \hat{x} + 2\hat{y} + 3\hat{z}, \vec{v} = (2m/s) \hat{y} + (1m/s) \hat{z} \quad (3.39)$$

Example 3: Consider the particle in the previous example. What is the distance of closest approach (the “impact parameter”) of the particle to the origin?

Solution: Let the distance of the closest approach be described by a point $(x', y', z')$

$$\vec{r}' = x' \hat{x} + y' \hat{y} + z' \hat{z} \quad (3.40)$$

The closest distance is determined by the magnitude of the vector normal to the trajectory of the particle. That means we must have $\vec{r}' \perp \vec{r}' - \vec{r}_0$ for the impact parameter. We recall that when two vectors are perpendicular their dot product is zero

$$\vec{r}' \cdot (\vec{r}' - \vec{r}_0) = 0 \quad (3.41)$$

Since $\vec{r}'$ describes a point on the trajectory of the particle given by

$$\vec{r}'(t) = \vec{r}_0 + \vec{v} t \quad (3.42)$$
3.4. EQUATION OF A PLANE

at which the particle becomes closest to the origin, we may write

\[ \vec{r}' = \vec{r}_0 + \vec{v}t' \Rightarrow \vec{r}' - \vec{r}_0 = \vec{v}t'. \]  (3.43)

Then

\[ (\vec{r}_0 + \vec{v}t') \cdot \vec{v}t' = 0 \]
\[ (\vec{r}_0 \cdot \vec{v}) t' + (\vec{v} \cdot \vec{v}) t'^2 = 0 \]  (3.44)

Noting that

\[ \vec{v} \cdot \vec{v} = (2m/s)^2 + (1m/s)^2 = 5m/s \]  (3.45)
\[ \vec{r}_0 \cdot \vec{v} = r_{0x}v_x + r_{0y}v_y + r_{0z}v_z = 2 \times 2 + 3 \times 1 = 7 \]

Therefore

\[ (\vec{r}_0 \cdot \vec{v}) t' + (\vec{v} \cdot \vec{v}) t'^2 = 0 \]  (3.46)

becomes

\[ 7t' + 5t'^2 = 0 \Rightarrow t' = 0, t' = \frac{7}{5} \]  (3.47)

The acceptable value is \( t' = 0 \). Thus the shortest distance from the origin is given by the magnitude of the vector

\[ \vec{r}' = \vec{r}_0 + \vec{v}t' = \vec{r}_0 \Rightarrow \vec{r}' = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14} \]  (3.48)

**Example 4** Find the shortest distance from the origin to the plane

\[ 3x - 2y - 6z = 7 \]  (3.49)

**Solution:** Comparing

\[ 3x - 2y - 6z = 7 \]  (3.50)

to the equation of a plane

\[ ax + by + cz = d \]  (3.51)

we have

\[ \vec{N} = 3\hat{x} - 2\hat{y} - 6\hat{z} \]  (3.52)

which is a vector normal to the plane. The unit vector along normal vector is given by

\[ \hat{N} = \frac{\vec{N}}{|\vec{N}|} = \frac{3\hat{x} - 2\hat{y} - 6\hat{z}}{\sqrt{3^2 + (-2)^2 + (-6)^2}} \Rightarrow \hat{N} = \frac{3}{\sqrt{14}}\hat{x} - \frac{2}{\sqrt{14}}\hat{y} - \frac{6}{\sqrt{14}}\hat{z} \]  (3.53)

We now pick any point on the plane. This point must satisfy the equation

\[ 3x - 2y - 6z = 7. \]  (3.54)
We may pick \((3, 1, 0)\) which we write as
\[
\vec{r} = 3\hat{x} + 1\hat{y}
\]  
(3.55)

Then the distance from the origin to the plane is the component of this vector along the vector perpendicular to the plane, \(\vec{N}\), because the distance is measured by the length of the normal to the plane. We can write this distance as
\[
d = \vec{r} \cdot \vec{N} = \frac{3 \times 3}{7} - \frac{2 \times 1}{7} \Rightarrow d = 1 \text{ units}
\]  
(3.56)

**Example 5** Find the symmetric and parametric equations of the line through \((0, -2, 4)\) and \((3, -2, 1)\).

**Solution:** The symmetric equations of a line are given by
\[
\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}
\]  
(3.57)

where
\[
\vec{A} = a\hat{x} + b\hat{y} + c\hat{z}
\]  
(3.58)

is a vector parallel to the line passing through \((0, -2, 4)\) and \((3, -2, 1)\).

Noting that a vector parallel to the line passing through can be expressed as
\[
\vec{A} = (3 - 0)\hat{x} + (-2 - -2)\hat{y} + (1 - 4)\hat{z} \Rightarrow \vec{A} = 3\hat{x} + 0\hat{y} - 3\hat{z}
\]  
(3.59)

we have
\[
a = 3, b = 0, c = -3
\]  
(3.60)

so that the symmetric equations becomes
\[
\frac{x - 0}{3} = -\frac{z - 4}{3}, y = -2
\]  
(3.61)

The parametric equations are given by
\[
\vec{r} = \vec{r}_0 + \vec{A}t \text{ or } \begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{cases}
\]  
(3.62)

which give
\[
\vec{r} = \begin{dcases} x = 3t \\ y = -2 \\ z = 4 - 3t \end{dcases}
\]  
(3.63)

**Example 6** Find the equation of the plane containing the three points \((0, 1, 1)\), \((2, 1, 3)\), and \((4, 2, 1)\).
3.5. HOMEWORK ASSIGNMENT 6

Solutions: We first construct two vectors $\vec{A}$ and $\vec{B}$ that lie on the plane formed by these three points. Consider the vector pointing from $(0, 1, 1)$ to $(2, 1, 3)$ and the vector from $(0, 1, 1)$ to $(4, 2, 1)$ and let denote these vectors by $\vec{A}$ and $\vec{B}$, respectively. The we can write

$$\vec{A} = (2 - 0) \hat{x} + (1 - 1) \hat{y} + (3 - 1) \hat{z} = 2\hat{x} + 2\hat{z} \quad (3.64)$$

$$\vec{B} = (4 - 0) \hat{x} + (4 - 1) \hat{y} + (1 - 1) \hat{z} = 4\hat{x} + 3\hat{y}$$

The vector, $\vec{N}$, that is normal to the plane formed by these two vectors (i.e. the three points) can be determined using the cross product of the two vectors

$$\vec{N} = \vec{A} \times \vec{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \hat{x} & \hat{y} & \hat{z} \\ 2 & 2 & 0 \end{vmatrix} = \hat{N} = \hat{x} \begin{vmatrix} 2 & 2 \\ 0 & 0 \end{vmatrix} - \hat{y} \begin{vmatrix} 0 & 2 \\ 4 & 0 \end{vmatrix} + \hat{z} \begin{vmatrix} 0 & 2 \\ 4 & 3 \end{vmatrix}$$

$$\vec{N} = -6\hat{x} + 8\hat{y} - 8\hat{z} \quad (3.66)$$

We recall that equation of a plane is given by

$$a (x - x_0) + b (y - y_0) + c (z - z_0) = 0 \quad (3.67)$$

where

$$\vec{N} = a\hat{x} + b\hat{y} + c\hat{z} \quad (3.68)$$

and

$$\vec{r}_0 = x_0\hat{x} + y_0\hat{y} - z_0\hat{z} \quad (3.69)$$

is a point on the plane as measured from the origin. Then we can write the equation of the plane formed by the three points as

$$-6 (x - x_0) + 8 (y - y_0) - 8 (z - z_0) = 0 \quad (3.70)$$

where

$$\vec{r}_0 = x_0\hat{x} + y_0\hat{y} - z_0\hat{z} \quad (3.71)$$

is given by $\vec{r}_0 = (0, 1, 1)$, $(2, 1, 3)$, or $(4, 2, 1)$.

3.5 Homework Assignment 6

1. Find the angles between $\vec{A} = -2i + j - 2k$ and $\vec{B} = 2i - 2j$

2. Show that the vectors $\vec{A} = 2i - j + 4k$ and $\vec{B} = 5i + 2j - 2k$ are orthogonal (perpendicular) and find a third vector perpendicular to both.

3. Show that the lines that join $(0, 0, 0)$ to $(1, 2, -1)$ and $(1, 1, 1)$ to $(2, 3, 4)$ intersect and find the acute angle between them.
4. A particle is traveling along the line \((x - 3)/2 = (y + 1)/(-2) = z - 1\). Write the equation of this path in the form \(\vec{r} = \vec{r}_0 + \vec{A}t\). Find the distance of closest approach of the particle to the origin (that is, the distance from the origin to the line). If \(t\) represents time, show that the time of the closest approach is

\[
t = -\frac{\vec{A} \cdot \vec{r}_0}{\vec{A}^2}.
\]

Use this value to check your answer for the distance of closest approach. If \(P\) is the point of closest approach, what is \(\vec{A} \cdot \vec{r}\)?

5. The vectors \(\vec{A} = ai + bj\) and \(\vec{B} = ci + dj\) form two sides of a parallelogram. Show that the area of the parallelogram is given by the absolute value of the following determinant

\[
\begin{vmatrix}
a & b \\
c & d
\end{vmatrix}
\]
4.1 Important Terminologies

A matrix: A rectangular array of quantities, usually inclosed in large parentheses with \( n \) rows and \( m \) columns.

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1m-1} & a_{1m} \\
a_{21} & a_{22} & \cdots & a_{2m-1} & a_{2m} \\
& \ddots & \ddots & \ddots & \vdots \\
& & \ddots & \ddots & \ddots \\
a_{n-11} & a_{n-12} & \cdots & a_{n-1m-1} & a_{n-1m} \\
a_{n1} & a_{n2} & \cdots & a_{nm-1} & a_{nm}
\end{bmatrix}
\]  \hspace{1cm} (4.1)

when \( n = m \) the matrix is called a square matrix.

Transpose of a Matrix: The transpose of a Matrix \( A \) in Eq. (4.1) which we denote by \( A^T \) is obtained by simply write the rows of matrix \( A \) as columns.

\[
A^T = \begin{bmatrix}
a_{11} & a_{21} & \cdots & a_{n-11} & a_{n1} \\
a_{12} & a_{22} & \cdots & a_{n-12} & a_{n2} \\
& \ddots & \ddots & \ddots & \vdots \\
& & \ddots & \ddots & \ddots \\
a_{1m-1} & a_{2m-1} & \cdots & a_{n-1m-1} & a_{nm-1} \\
a_{1m} & a_{2m} & \cdots & a_{n-1m} & a_{nm}
\end{bmatrix}
\]  \hspace{1cm} (4.2)

\[
A = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix}, A^T = \begin{pmatrix} 2 & 2 \\ 3 & 1 \end{pmatrix} \hspace{1cm} (4.3)
\]

\[
B = \begin{pmatrix} 2 & 4 \\ 1 & -1 \\ 3 & -1 \end{pmatrix}, B^T = \begin{pmatrix} 2 & 1 & 3 \\ 4 & -1 & -1 \end{pmatrix}
\]
The Identity Matrix; \( \mathbf{I} \) (The Unit Matrix, \( \mathbf{U} \))

\[ \mathbf{I} \mathbf{A} = \mathbf{A} \mathbf{I} = \mathbf{A} \quad (4.4) \]

### 4.2 Matrix Arithmetic and Manipulation

Consider the three matrices:

\[
\mathbf{A} = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 4 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & -1 & -2 \\ -1 & 0 & 1 \end{pmatrix}, \quad (4.5)
\]

\[
\mathbf{D} = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 2 \\ 3 & 1 & 0 \end{pmatrix}
\]

**Multiplication by a Scalar**: Any matrix can be multiplied by a scalar:

\[
2\mathbf{A} = \begin{pmatrix} 2 \times 2 & 3 \times 2 & 1 \times 2 & -4 \times 2 \\ 2 \times 2 & 1 \times 2 & 0 \times 2 & 5 \times 2 \end{pmatrix} = \begin{pmatrix} 4 & 6 & 2 & -8 \\ 4 & 2 & 0 & 10 \end{pmatrix} \quad (4.6)
\]

**Addition and subtraction**: Two matrices can be added or subtracted if and only if they have the same dimensions. From matrices \( \mathbf{A}, \mathbf{B}, \mathbf{C}, \) and \( \mathbf{D} \) we can add/subtract only matrices \( \mathbf{C} \) and \( \mathbf{D} \)

\[
\mathbf{C} + \mathbf{D} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & -1 & -2 \\ -1 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -2 & 0 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 4 \\ 5 & -2 & 0 \\ 2 & 1 & 1 \end{pmatrix} \quad (4.7)
\]

**Matrix Multiplication**: two matrices can be multiplied if and only if the number of columns of the first matrix is equal to the number of rows of the second matrix. If matrices have the same dimension, then they can be multiplied. From the above matrices we can make the multiplications:

\[
\mathbf{AB} = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 1 & -1 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} (ab)_{11} & (ab)_{12} \\ (ab)_{21} & (ab)_{22} \end{pmatrix} \quad (4.8)
\]

\[
\mathbf{CD} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & -1 & -2 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & 0 \end{pmatrix} = \begin{pmatrix} (cd)_{11} & (cd)_{12} & (cd)_{13} \\ (cd)_{21} & (cd)_{22} & (cd)_{23} \\ (cd)_{31} & (cd)_{32} & (cd)_{33} \end{pmatrix} \quad (4.9)
\]

but we can not make the matrix multiplications \( \mathbf{BC} \) or \( \mathbf{BD} \)
4.3. MATRIX REPRESENTATION OF A SET OF LINEAR EQUATIONS

The element in row \( i \) and column \( j \) of the product matrix \( AB \) is equal to row \( i \) of \( A \) times column \( j \) of \( B \). In index notation

\[
(ab)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}, \tag{4.10}
\]

where \((ab)_{ij}\) is the element of the product matrix \(AB\).

Commutativities: For any two multiplyable matrices \(C\) and \(D\),

\[
CD \neq DC \tag{4.11}
\]

Commutator: For square matrices \(C\) and \(D\) the Commutator \([C, D]\) is defined as

\[
[C, D] = CD - DC \tag{4.12}
\]

For any three matrices, \(F, G,\) and \(H\) that can be multiplied we can write

The Associative Law:

\[
F(GH) = (FG)H \tag{4.13}
\]

The Distributive Law:

\[
F(G + H) = FG + FH \tag{4.14}
\]

4.3 Matrix representation of a set of linear equations

A set of linear equations

\[
a_1x + b_1y + c_1z + d_1w = c_1, \tag{4.15}
a_2x + c_2z + d_2w = c_2,
b_3y + c_3z = c_4,
a_4x = c_4,
\]

can be expressed using matrices as

\[
\begin{bmatrix}
a_1 & b_1 & c_1 & d_1 \\
a_2 & 0 & c_2 & d_2 \\
0 & b_3 & c_3 & 0 \\
a_4 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
w
\end{bmatrix}
=
\begin{bmatrix}
c_1 \\
c_2 \\
c_3 \\
c_4
\end{bmatrix}. \tag{4.16}
\]

or

\[
Mr = k, \tag{4.17}
\]

where

\[
M = \begin{bmatrix}
a_1 & b_1 & c_1 & d_1 \\
a_2 & 0 & c_2 & d_2 \\
0 & b_3 & c_3 & 0 \\
a_4 & 0 & 0 & 0
\end{bmatrix} \tag{4.18}
\]
is called the coefficient matrix,

\[
r = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix},
\]

and

\[
k = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}.
\]

One can augment the matrices \( M \) and \( k \) and write a matrix

\[
A = \begin{bmatrix} a_1 & b_1 & c_1 & d_1 & c_1 \\ a_2 & 0 & c_2 & d_2 & c_2 \\ 0 & b_3 & c_3 & 0 & c_3 \\ a_4 & 0 & 0 & 0 & c_4 \end{bmatrix}
\]

which is referred as the Augmented Matrix.

### 4.4 Matrix application: solving a set of linear equations

With application of matrices we can solve a set of linear equations. There are two different methods. The first involves the Gaussian Elimination method and Row Echelon Form whereas the second method involves Cramer’s Rule. We first discuss the first method. The second method will be discussed after we get introduced to square matrices and determinants.

**Gaussian Elimination method** (named after German mathematician and scientist Carl Friedrich Gauss): in this method we follow the following basic steps to solve a set of linear equations:

1. Write the set of linear equations using matrices.
2. Obtain the augmented matrix.
3. Apply any of the following elementary row operations
   
   (a) Multiply any row of the matrix by a non-zero constant,
   
   (b) Interchange any two rows,
   
   (c) Replace any row by a linear combination of the rows of the matrix (as long as the linear combination includes the row being replaced). In other words you can add to or subtract from a multiple of one row to another,
4.4. **MATRIX APPLICATION: SOLVING A SET OF LINEAR EQUATIONS**

4\textsuperscript{th} Repeat step 2 until the augmented matrix is in *Row Echelon Form* and entirely row reduced.

*Row Echelon Form*: a matrix is in *Row Echelon Form* if all of the following conditions are satisfied

- All nonzero rows (rows with at least one nonzero element) are above any rows of all zeroes.
  
  *Note*: All zero rows, if any, belong at the bottom of the matrix.

- The leading coefficient (the first nonzero number from the left, also called the pivot) of a nonzero row is always strictly to the right of the leading coefficient of the row above it.

- All entries in a column below a leading entry are zeroes (implied by the first two criteria).

*Reduced row echelon form (row canonical form)*: A matrix that is in row echelon form and also its every leading coefficient is 1.

**Example 1** The matrix

\[
B = \begin{bmatrix}
0 & 1 & 4 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]  

(4.22)

in *Row Echelon Form* and in a reduced row form where as the matrix

\[
C = \begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 9 & 0 & 2 \\
0 & 0 & 0 & 3
\end{bmatrix}
\]  

(4.23)

is in *Row Echelon Form* but not in reduced row form. The following matrix is not in *Row Echelon Form*

\[
C = \begin{bmatrix}
1 & 2 & 3 & 4 \\
0 & 3 & 7 & 2 \\
0 & 2 & 0 & 0
\end{bmatrix}
\]  

(4.24)

as the leading coefficient of row 3 is not strictly to the right of the leading coefficient of row 2.

*Rank of a Matrix*: the number of nonzero rows remaining when a matrix has been *row reduced* is called the rank of the matrix.

(a) If \((\text{rank } M) < (\text{rank } A)\), the equations are inconsistent and there is no solution.

(b) If \((\text{rank } M) = (\text{rank } A) = n\) (number of unknowns), there is one solution.
Example 2 Consider the circuit shown. The circuit element values are $R = 1.0\, \text{W}$ and $\Delta V = 2.0\, \text{V}$.

(a) Applying Kirchhoff’s voltage and current rules find set of linear equations.

(b) Find the coefficient and augmented matrices for the set of linear equations.

(c) Find the values of the three branch currents $I_1$, $I_2$, and $I_3$ using Gaussian elimination method.

(d) Find the rank of the coefficient and augmented matrices. Is the ranks consistent with what we discussed earlier?

**Kirchhoff’s Voltage Law:** Start at one point in a circuit and go around any closed loop in the circuit. The voltage at the starting point of the loop must be the same as the voltage at the ending point, since they are the same point in the circuit. Another way of stating this is simply that the net change in voltage as you go around any closed loop in the circuit must be zero (so the starting potential must be the same as the ending potential). This is just a restatement of the conservation of energy.

**Kirchhoff’s Current Law:** The net current flowing into any node of a circuit must equal the net current flowing out of that same node. This law says that there is no place from which charge magically appears in the circuit, and no place that the charge disappears. There is a net conservation of charge in an electric circuit.

**Solution:**

(a) Applying Kirchhoff’s voltage law for the closed loop $abe$ we can write

\[
\Delta V + I_1R + 2I_2R - 2\Delta V + I_1R = 0 \Rightarrow 2I_1R + 2I_2R = \Delta V \Rightarrow I_1 + I_2 = \frac{\Delta V}{2R}
\]

and for closed loop $bcde$

\[
I_3R - 2\Delta V + I_3R + 2\Delta V - 2I_2R = 0 \Rightarrow 2I_3R - 2I_2R = 0 - I_2 + I_3 = 0.
\]
4.4. MATRIX APPLICATION: SOLVING A SET OF LINEAR EQUATIONS

Using Kirchhoff’s Current Law at node b, we may write

\[ I_1 = I_2 + I_3 \Rightarrow I_1 - I_2 - I_3 = 0 \]  

(4.27)

(b) Substituting \( R = 1 \) and \( \Delta V = 2V \) in the results we obtained above, we have

\[
\begin{align*}
I_1 + I_2 &= 1 \\
-I_2 + I_3 &= 0 \\
I_1 - I_2 - I_3 &= 0
\end{align*}
\]

or

\[
\begin{align*}
I_1 + I_2 + 0I_3 &= 1 \\
0I_1 - I_2 + I_3 &= 0 \\
I_1 - I_2 - I_3 &= 0
\end{align*}
\]

so that we can put these equations using matrices as

\[ MI = K, \]  

(4.30)

where coefficient matrix \( M \) is given by

\[
M = \begin{bmatrix}
1 & 1 & 0 \\
0 & -1 & 1 \\
1 & -1 & -1
\end{bmatrix}
\]  

(4.31)

and

\[
I = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}, K = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.
\]  

(4.32)

The augmented matrix is a combination of the coefficient and the constant matrix given by

\[
A = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & -1 & 1 & 0 \\
1 & -1 & -1 & 0
\end{bmatrix}
\]  

(4.33)

(c) We apply elementary row operations to row reduce the augmented matrix.

\[
A = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & -1 & 1 & 0 \\
1 & -1 & -1 & 0
\end{bmatrix}
\]  

(4.34)

Subtract row 1 from row 3:

\[
A = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & -1 & 1 & 0 \\
0 & -2 & -1 & -1
\end{bmatrix}
\]  

(4.35)
Multiply row 2 by 2 and subtract from row three:

\[
A = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & -1 & 1 & 0 \\
0 & 0 & -3 & -1
\end{bmatrix}
\]  \hspace{1cm} (4.36)

Divide row 3 by \(-3\):

\[
A = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & -1 & 1 & 0 \\
0 & 0 & 1 & \frac{1}{3}
\end{bmatrix}
\]  \hspace{1cm} (4.37)

Subtract row 3 from row 2:

\[
A = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & -1 & 0 & -\frac{1}{3} \\
0 & 0 & 1 & \frac{1}{3}
\end{bmatrix}
\]  \hspace{1cm} (4.38)

Divide row 2 by \(-1\):

\[
A = \begin{bmatrix}
1 & 1 & 0 & 1 \\
0 & 1 & 0 & \frac{1}{3} \\
0 & 0 & 1 & \frac{1}{3}
\end{bmatrix}
\]  \hspace{1cm} (4.39)

Subtract row 2 from row 1:

\[
A = \begin{bmatrix}
1 & 0 & 0 & \frac{2}{3} \\
0 & 1 & 0 & \frac{1}{3} \\
0 & 0 & 1 & \frac{1}{3}
\end{bmatrix}
\]  \hspace{1cm} (4.40)

Therefore the values of the three currents would be

\[
\begin{bmatrix}
I_1 \\
I_2 \\
I_3
\end{bmatrix} = \begin{bmatrix}
\frac{2}{3} \\
\frac{1}{3} \\
\frac{1}{3}
\end{bmatrix}
\]  \hspace{1cm} (4.41)

which means

\[
I_1 = \frac{2}{3}A, I_2 = \frac{1}{3}A, I_3 = \frac{1}{3}A
\]  \hspace{1cm} (4.42)

(d) The number of nonzero rows for the row reduced augmented matrix

\[
A = \begin{bmatrix}
1 & 0 & 0 & \frac{2}{3} \\
0 & 1 & 0 & \frac{1}{3} \\
0 & 0 & 1 & \frac{1}{3}
\end{bmatrix}
\]  \hspace{1cm} (4.43)

is 3 and the rank is 3. For the reduced coefficient matrix (the first three columns in the reduced augmented matrix)

\[
M = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]  \hspace{1cm} (4.44)

we have also 3 nonzero rows and the rank is 3. There are 3 unknowns, this means that \(\text{rank } M = \text{rank } A = 3\) (number of unknowns), therefore we find one solution for each unknowns.
4.5 Determinant of a square matrix

The determinant of a $2 \times 2$ matrix: Suppose we are given a $2 \times 2$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

the determinant is given by

$$|A| = \sum_{j=1}^{2} (-1)^{i+j} a_{ij} M_{ij},$$

where $M_{ij}$ is called the minor of $a_{ij}$. The minor of $a_{ij}$, $M_{ij}$, is the determinant of the matrix formed by removing the row and the column containing $a_{ij}$.

If we chose $i = 1$, we have

$$|A| = \sum_{j=1}^{2} (-1)^{1+j} a_{1j} M_{1j} = (-1)^{1+1} a_{11} M_{11} + (-1)^{1+2} a_{12} M_{12}$$

so that noting that

$$M_{11} = a_{22}, M_{12} = a_{21}$$

one finds

$$|A| = a_{11} a_{22} - a_{12} a_{21}. \quad (4.49)$$

The determinant of a $3 \times 3$ matrix: Consider a 3-dimensional matrix

$$C = \begin{bmatrix} i & j & k \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{bmatrix} \quad (4.50)$$

constructed from two 3-dimensional vectors in Cartesian coordinates and the unit vectors, $i$, $j$, and $k$. Using Eq. (4.59) we may write the determinant of this matrix as

$$|A| = \sum_{j=1}^{3} (-1)^{1+j} a_{1j} M_{ij}$$

$$= (-1)^{1+1} a_{11} M_{11} + (-1)^{1+2} a_{12} M_{12} + (-1)^{1+3} a_{13} M_{13},$$

where we chose $i = 1$. Note that for a 3-dimensional square matrix $i = 1, 2, \text{ or } 3$.

Noting that

$$a_{11} = i, a_{12} = j, \text{ and } a_{13} = k$$

the corresponding Minors $M_{11}, M_{12}, \text{ and } M_{13}$ are given by

$$M_{11} = \begin{vmatrix} A_y & A_z \\ B_y & B_z \end{vmatrix} = A_y B_z - A_z B_y, \quad (4.53)$$

$$M_{12} = \begin{vmatrix} A_x & A_z \\ B_x & B_z \end{vmatrix} = A_x B_z - A_z B_x, \quad (4.54)$$

$$M_{13} = \begin{vmatrix} A_x & A_y \\ B_x & B_y \end{vmatrix} = A_x B_y - A_y B_x, \quad (4.55)$$
we find for Eq. (4.51)

\[ |C| = (-1)^{1+i+1} (A_y B_z - A_z B_y) + (-1)^{1+j+2} (A_x B_z - A_z B_x) + (-1)^{1+k+3} (A_x B_y - A_y B_x) \]

\[ \Rightarrow |C| = i (A_y B_z - A_z B_y) + j (A_x B_z - A_z B_x) + k (A_x B_y - A_y B_z) \]

Note: The result in Eq. (4.56) shows that the determinant of the matrix \( C \) is the cross product of vector \( \vec{A} \) and \( \vec{B} \).

\[ \vec{A} \times \vec{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \]  

(4.57)

We now consider a square, \( n \times n \), matrix \( A \) (a matrix with \( n \) rows and \( n \) columns)

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n-1} & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n-1} & a_{2n} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in-1} & a_{in} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
a_{n-11} & a_{n-12} & \cdots & a_{n-1j} & \cdots & a_{n-1n-1} & a_{n-1n} \\
a_{n1} & a_{n2} & a_{nj} & a_{nn} \\
\end{bmatrix}
\]  

(4.58)

Then the determinant of this matrix (represented as \( |A| \) or \( \det A \)) is given by

\[ |A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \]  

(4.59)

where \( M_{ij} \) is called the minor of \( a_{ij} \). The minor of \( a_{ij} \), \( M_{ij} \), is the determinant of the matrix formed by removing the row and the column containing \( a_{ij} \).

**Useful properties of determinants:** It is useful to know the following properties of determinants of a square matrix which can be proved using Eq. (4.59).

1. **Multiplying by a constant:** Suppose a row or a column of a square matrix \( A \) is multiplied by a constant \( k \), then the determinant of the new matrix is \( k \) times the determinant of the matrix \( A \).

**Example 3** For the 2-dimensiona square matrix

\[
A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}
\]  

(4.60)

we may write a new matrix, \( B \) given by

\[
B = \begin{bmatrix} ka_{11} & a_{12} \\ ka_{21} & a_{22} \end{bmatrix} \quad \text{or} \quad B = \begin{bmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{bmatrix}
\]  

(4.61)
4.5. **DETERMINANT OF A SQUARE MATRIX**

then

\[ \det B = k \det A, \]  \hspace{1cm} (4.62)

2. **Zero determinant value:** The value of the determinant of square matrix is zero if

(a) all elements of a row (or a column) are zero

**Example 4**

\[
A = \begin{bmatrix} 0 & a_{12} \\ 0 & a_{22} \end{bmatrix} \text{ or } A = \begin{bmatrix} a_{11} & a_{12} \\ 0 & 0 \end{bmatrix} \Rightarrow \det A = 0, \hspace{1cm} (4.63)
\]

(b) two rows (or columns) are identical

**Example 5**

\[
A = \begin{bmatrix} a_{11} & a_{12} = a_{11} \\ a_{21} & a_{22} = a_{21} \end{bmatrix} \text{ or } A = \begin{bmatrix} a_{11} & a_{21} \\ a_{22} = a_{21} & a_{22} = a_{22} \end{bmatrix} \Rightarrow \det A = 0
\]

(4.64)

(c) two rows (or columns) are proportional

**Example 6**

\[
B = \begin{bmatrix} a_{11} & a_{12} = ka_{11} \\ a_{21} & a_{22} = ka_{21} \end{bmatrix} \text{ or } B = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} = ka_{21} & a_{22} = ka_{22} \end{bmatrix} \Rightarrow \det A = 0
\]

(4.65)

where \( k \) is the proportionality constant.

3. **Interchanging a row or a column:** If two rows (or two columns) of a square matrix are interchanged, the determinant of the matrix remain the same.

**Example 7**

\[
A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \text{ or } B = \begin{bmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{bmatrix} \Rightarrow \det A = \det B
\]

(4.66)

**Example 8** Evaluate the determinant of the matrix \( A \):

\[
A = \begin{pmatrix} 2 & 2 & 0 \\ 0 & -2 & 2 \\ 1 & -1 & -1 \end{pmatrix}
\]

(4.67)

**Solution:** Applying the relation above we can express the determinant as

\[
|A| = 2(-1)^{1+1} \begin{pmatrix} -2 & 2 \\ -1 & -1 \end{pmatrix} + 2(-1)^{1+2} \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix}
\]

(4.68)

\[ + 0(-1)^{1+3} \begin{pmatrix} 2 & 2 & 0 \\ 0 & -2 & 2 \\ 1 & -1 & -1 \end{pmatrix} \]

(4.69)
\[ |A| = 2 \begin{pmatrix} -2 & 2 \\ -1 & -1 \end{pmatrix} - 2 \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix} \]  
\[
|A| = 2((-2 \times -1) - (2 \times -1)) - 2((0 \times -1) - (2 \times 1)) = 8 + 4 \\
\Rightarrow \quad |A| = 12
\]

4.6 Cramer’s Rule

Let a system of simultaneous equations be represented by the augmented matrix \( A \), and let \( M \) be the corresponding coefficient matrix. Also, let \( M_j \) be the coefficient matrix with its \( j^{th} \) column replaced by the constant column from \( A_{aug} \). We define the following determinants:

\[ D = |M|, D_j = |M_j| \]  

The value of the \( j^{th} \) unknown, \( x_j \), is then equal to

\[ x_j = \frac{D_j}{D} \]  

provided \( D \neq 0 \)

**Example 9** Solve the system of equations from Example 2 using Cramer’s rule.  
( \( D = 12, D_1 = 8, D_2 = 4, \) and \( D_3 = 4. \) )

**Solution:** We recall that the coefficient matrix is given by

\[ M = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \]  

and the corresponding Augmented matrix by

\[ A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{bmatrix}. \]  

Replacing the first column of \( M \) by the last column of \( A \), we find

\[ M_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & -1 \end{bmatrix}, \]  

the second column of \( M \) by the last column of \( A \)

\[ M_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \]  

\[ M_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \]
and the third column of $M$ by the last column of $A$

$$M_3 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & -1 & 0 \end{bmatrix}. \tag{4.78}$$

For the determinants of the matrices $M$, $M_1$, $M_2$, and $M_3$, one can easily find

$$|M| = \begin{vmatrix} -1 & 1 \\ -1 & -1 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} \Rightarrow |M| = 2 + 1 = 3, \tag{4.79}$$

$$|M_1| = \begin{vmatrix} -1 & 1 \\ -1 & -1 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 0 & -1 \end{vmatrix} \Rightarrow |M_1| = 2, \tag{4.80}$$

$$|M_2| = \begin{vmatrix} 0 & 1 \\ 0 & -1 \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} \Rightarrow |M_2| = 1, \tag{4.81}$$

and

$$|M_3| = \begin{vmatrix} -1 & 0 \\ -1 & 0 \end{vmatrix} - \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} + \begin{vmatrix} 0 & -1 \\ 1 & -1 \end{vmatrix} \Rightarrow |M_3| = 1. \tag{4.82}$$

Therefore, the three currents are given by

$$I_1 = \frac{|M_1|}{|M|} = \frac{2}{3}, I_2 = \frac{|M_2|}{|M|} = \frac{1}{3}, I_2 = \frac{|M_3|}{|M|} = \frac{1}{3}. \tag{4.83}$$

### 4.7 The adjoint and inverse of a matrix

**Adjoint of a Matrix**: The adjoint of a square matrix, $A$, is given by

$$\text{adj}(A) = [\text{cof}(A)]^T \tag{4.84}$$

where $\text{cof}(A)$ is the cofactor of the matrix $A$. We recall that the minor of matrix $A$ ($M_{ij}$) is the determinant of the matrix formed from matrix $A$ by removing the $i^{th}$ row and $j^{th}$ column. For the cofactor matrix the elements are expressed as

$$[\text{cof}(A)]_{ij} = (-1)^{i+j} M_{ij}. \tag{4.85}$$

**Inverse of a (Square) Matrix**: $A^{-1}$

$$A^{-1} A = AA^{-1} = I \tag{4.86}$$

We can determine the inverse of an invertible matrix ($\det A \neq 0$) using row reduction or the adjoint matrix.

**a. Row reduction** in this approach for the matrix, for example,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{4.87}$$
we start from
\[
\begin{bmatrix}
a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\
a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\
a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \\
\end{bmatrix}
\]
and do elementary row operation until we end up with
\[
\begin{bmatrix}
1 & 0 & 0 & b_{11} & b_{12} & b_{13} \\
0 & 1 & 0 & b_{21} & b_{22} & b_{23} \\
0 & 0 & 1 & b_{31} & b_{32} & b_{33} \\
\end{bmatrix}
\]
so that the inverse of the Matrix \( A \) is given by
\[
A^{-1} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}
\]

b. Using the adjoint matrix: Using the adjoint matrix the inverse can be expressed as
\[
A^{-1} = \frac{[\text{cof}(A)]^T}{\det |A|}
\]

Example 10 Find the inverse of the matrix
\[
A = \begin{pmatrix}
-1 & 2 & 3 \\
2 & 0 & -4 \\
1 & -1 & 1 \\
\end{pmatrix}
\]
using

a. Row reduction approach

b. The adjoint matrix approach

Solution: a. In the row reduction approach we start from
\[
\begin{bmatrix}
-1 & 2 & 3 & 1 & 0 & 0 \\
2 & 0 & -4 & 0 & 1 & 0 \\
1 & -1 & 1 & 0 & 0 & 1 \\
\end{bmatrix}
\]
and try to get
\[
\begin{bmatrix}
1 & 0 & 0 & a_{11} & a_{12} & a_{13} \\
0 & 1 & 0 & a_{21} & a_{22} & a_{23} \\
0 & 0 & 1 & a_{31} & a_{32} & a_{33} \\
\end{bmatrix}
\]
so that we can get the inverse matrix
\[
A^{-1} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}
\]
In order to get this matrix let's make the elementary row operation.

*Add row 1 to row 3:*

\[
\begin{bmatrix}
-1 & 2 & 3 & 1 & 0 & 0 \\
2 & 0 & -4 & 0 & 1 & 0 \\
0 & 1 & 4 & 1 & 0 & 1
\end{bmatrix}
\]  
(4.95)

*Add row 3 to row 2:*

\[
\begin{bmatrix}
-1 & 2 & 3 & 1 & 0 & 0 \\
2 & 1 & 0 & 1 & 1 & 1 \\
0 & 1 & 4 & 1 & 0 & 1
\end{bmatrix}
\]  
(4.96)

*Multiply row 2 by 2 and subtract the result from row 1*

\[
\begin{bmatrix}
-5 & 0 & 3 & -1 & -2 & -2 \\
2 & 1 & 0 & 1 & 1 & 1 \\
0 & 1 & 4 & 1 & 0 & 1
\end{bmatrix}
\]  
(4.97)

*Divide row 1 by -5 and row 2 by 2:*

\[
\begin{bmatrix}
1 & 0 & -3/5 & 1/5 & 2/5 & 2/5 \\
1/2 & 0 & 1/2 & 1/2 & 1/2 \\
0 & 1/4 & 1 & 0 & 1
\end{bmatrix}
\]  
(4.98)

*Subtract row 1 from row 2*

\[
\begin{bmatrix}
1 & 0 & -3/5 & 1/5 & 2/5 & 2/5 \\
0 & 1/2 & 3/5 & 3/10 & 1/10 & 1/10 \\
0 & 1/4 & 1 & 0 & 1
\end{bmatrix}
\]  
(4.99)

*Divide row 2 by 1/2*

\[
\begin{bmatrix}
1 & 0 & -3/5 & 1/5 & 2/5 & 2/5 \\
0 & 1 & 6/5 & 3/5 & 1/5 & 1/5 \\
0 & 1 & 4 & 1 & 0 & 1
\end{bmatrix}
\]  
(4.100)

*Subtract row 2 from row 3*

\[
\begin{bmatrix}
1 & 0 & -3/5 & 1/5 & 2/5 & 2/5 \\
0 & 1 & 6/5 & 3/5 & 1/5 & 1/5 \\
0 & 0 & 14/5 & 2/5 & -1/5 & 4/5
\end{bmatrix}
\]  
(4.101)

*Divide row 1 by -3/5, row 2 by 6/5, and row 3 by 14/5*

\[
\begin{bmatrix}
-5/3 & 0 & 1 & -1/3 & -2/3 & -2/3 \\
0 & 5/6 & 1 & 1/2 & 1/6 & 1/6 \\
0 & 0 & 1 & 1/7 & -1/14 & 2/7
\end{bmatrix}
\]  
(4.102)
subtract row 3 from row 2
\[
\begin{bmatrix}
-5/3 & 0 & 1 \\
0 & 5/6 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
-1/3 & -2/3 & -2/3 \\
5/14 & 5/21 & 5/42 \\
1/7 & -1/14 & 2/7
\end{bmatrix}
\]
(4.103)

subtract row 3 from row 1
\[
\begin{bmatrix}
-5/3 & 0 & 0 \\
0 & 5/6 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
-10/21 & -25/42 & -20/21 \\
5/14 & 5/21 & 5/42 \\
1/7 & -1/14 & 2/7
\end{bmatrix}
\]
(4.104)

Divide row 1 by \(-5/3\) and row 2 by 5/6
\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
1/7 & 5/14 & 4/7 \\
3/7 & 2/7 & -1/7 \\
1/7 & -1/14 & 2/7
\end{bmatrix}
\]
(4.105)

Then the inverse matrix is given by
\[
A^{-1} = \begin{bmatrix}
2/7 & 5/14 & 4/7 \\
3/7 & 2/7 & -1/7 \\
1/7 & -1/14 & 2/7
\end{bmatrix}.
\]
(4.106)

Using Matemativa
\[
\text{In}[1]= \begin{bmatrix}
-1 & 2 & 3 \\
2 & 0 & -4 \\
1 & -1 & 1
\end{bmatrix}
\]
\[
\text{Out}[1]= \begin{bmatrix}
\frac{1}{7} & \frac{5}{14} & \frac{4}{7} \\
\frac{2}{7} & \frac{2}{7} & -\frac{1}{7} \\
\frac{1}{7} & -\frac{1}{14} & \frac{2}{7}
\end{bmatrix}
\]

N.B. Use Mathematica only to check your result!!!

b. Here we first need to find the adjoint matrix of \(A\). In order to find the adjoint, first we need to find the cofactor matrix. Let this cofactor matrix be
\[
cof (A) = \begin{bmatrix}
c_{11} & c_{12} & c_{13} \\
c_{21} & c_{22} & c_{23} \\
c_{31} & c_{32} & c_{33}
\end{bmatrix}
\]
(4.107)

where the elements
\[
c_{ij} = (-1)^{i+j} M_{ij}
\]
(4.108)

in which \(M_{ij}\) is the minor for the matrix \(A\). Now using
\[
A = \begin{bmatrix}
-1 & 2 & 3 \\
2 & 0 & -4 \\
1 & -1 & 1
\end{bmatrix}
\]
(4.109)
we can write

\[
M_{11} = \det \begin{vmatrix} 0 & -4 \\ -1 & 1 \end{vmatrix} = -4 \Rightarrow c_{11} = (-1)^2 M_{11} = -4 \quad (4.110)
\]

\[
M_{12} = \det \begin{vmatrix} 2 & -4 \\ 1 & 1 \end{vmatrix} = 6 \Rightarrow c_{12} = (-1)^3 M_{12} = -6 \quad (4.111)
\]

\[
M_{13} = \det \begin{vmatrix} 2 & 0 \\ 1 & -1 \end{vmatrix} = -2 \Rightarrow c_{13} = (-1)^4 M_{13} = -2 \quad (4.112)
\]

\[
M_{21} = \det \begin{vmatrix} 2 & 3 \\ -1 & 1 \end{vmatrix} = 5 \Rightarrow c_{21} = (-1)^3 M_{21} = -5 \quad (4.113)
\]

\[
M_{22} = \det \begin{vmatrix} -1 & 3 \\ 1 & 1 \end{vmatrix} = -4 \Rightarrow c_{22} = (-1)^4 M_{22} = -4 \quad (4.114)
\]

\[
M_{23} = \det \begin{vmatrix} -1 & 2 \\ 1 & -1 \end{vmatrix} = -1 \Rightarrow c_{23} = (-1)^5 M_{23} = 1 \quad (4.115)
\]

\[
M_{31} = \det \begin{vmatrix} 2 & 3 \\ 0 & -4 \end{vmatrix} = -8 \Rightarrow c_{31} = (-1)^4 M_{31} = -8 \quad (4.116)
\]

\[
M_{32} = \det \begin{vmatrix} -1 & 3 \\ 2 & -4 \end{vmatrix} = -2 \Rightarrow c_{32} = (-1)^5 M_{32} = 2 \quad (4.117)
\]

\[
M_{33} = \det \begin{vmatrix} -1 & 2 \\ 2 & 0 \end{vmatrix} = -4 \Rightarrow c_{33} = (-1)^6 M_{33} = -4 \quad (4.118)
\]

There follows that

\[
det(A) = \begin{bmatrix} -4 & -6 & -2 \\ -5 & -4 & 1 \\ -8 & 2 & -4 \end{bmatrix}
\]

and

\[
[ \text{det}(A) ]^T = \begin{bmatrix} -4 & -5 & -8 \\ -6 & -4 & 2 \\ -2 & 1 & -4 \end{bmatrix}.
\]

The determinant of \( A \) is given by

\[
\det |A| = \begin{vmatrix} -1 & 2 & 3 \\ 2 & 0 & -4 \\ 1 & -1 & 1 \end{vmatrix} = - \begin{vmatrix} 0 & -4 \\ -1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 2 & -4 \\ 1 & 1 \end{vmatrix} + 3 \begin{vmatrix} 2 & 0 \\ 1 & -1 \end{vmatrix} \quad (4.121)
\]

\[
\Rightarrow \ \det |A| = -14
\]

Then the inverse matrix

\[
A^{-1} = \frac{[ \text{det}(A) ]^T}{\det |A|}
\]

(4.122)
becomes
\[
A^{-1} = -\frac{1}{14} \begin{bmatrix}
-4 & -5 & -8 \\
-6 & -4 & 2 \\
-2 & 1 & -4
\end{bmatrix}
= \begin{bmatrix}
\frac{1}{7} & \frac{5}{14} & \frac{4}{7} \\
\frac{3}{7} & \frac{2}{7} & -\frac{1}{7} \\
\frac{1}{7} & -\frac{1}{14} & \frac{2}{7}
\end{bmatrix}
\] (4.123)

### 4.8 Orthogonal matrices and the rotation matrix

**Orthogonal Matrices**: matrices that make an orthogonal transformation of vectors. In an orthogonal transformation of vectors the magnitude of the vectors remains the same. For an orthogonal matrix

\[
M^{-1} = M^T
\] (4.124)

**The Rotation Operator**: Consider a rectangular coordinate system (XYZ) rotated by an angle \(\theta\) about the z axis in the counterclockwise direction to give another coordinate system (X0Y0Z0) as shown in the figure below.

Consider a vector \(\vec{r}\) which has components \((x, y, z)\) in the XYZ coordinate system. We want to find the components of this vector in the X'Y'Z' coordinate system. From the figure above we can write that the new components \((x', y', z')\) can be expressed as

\[
x' = x \cos (\theta) + y \cos (90 - \theta) = x \cos (\theta) + y \sin (\theta)
\]

\[
y' = -x \sin (\theta) + y \cos (\theta)
\]

\[
z' = z
\] (4.125)
4.8. ORTHOGONAL MATRICES AND THE ROTATION MATRIX

Using matrix representation we can write that
\[
\begin{pmatrix}
  x' \\
  y' \\
  z'
\end{pmatrix} =
\begin{pmatrix}
  \cos(\theta) & \sin(\theta) & 0 \\
  -\sin(\theta) & \cos(\theta) & 0 \\
  0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  x \\
  y \\
  z
\end{pmatrix}
\Rightarrow r' = Rr
\]

(4.126)

where the matrix \( R(\theta) = R \) is the rotation matrix given by
\[
R =
\begin{pmatrix}
  \cos(\theta) & \sin(\theta) & 0 \\
  -\sin(\theta) & \cos(\theta) & 0 \\
  0 & 0 & 1
\end{pmatrix}
\]

(4.127)

The inverse of the matrix \( R \) (I used Mathematica)

\[
\text{ln[6]} = \text{Simplify}[\text{Inverse}[\begin{pmatrix}
  \cos(\theta) & \sin(\theta) & 0 \\
  -\sin(\theta) & \cos(\theta) & 0 \\
  0 & 0 & 1
\end{pmatrix}] // \text{MatrixForm}
\]

\[
\text{Out[5]/MatrixForm} =
\begin{pmatrix}
  \cos(\theta) & -\sin(\theta) & 0 \\
  \sin(\theta) & \cos(\theta) & 0 \\
  0 & 0 & 1
\end{pmatrix}
\]

which is the same as the transpose of the matrix, \( R^T \) given by
\[
R^T =
\begin{pmatrix}
  \cos(\theta) & -\sin(\theta) & 0 \\
  \sin(\theta) & \cos(\theta) & 0 \\
  0 & 0 & 1
\end{pmatrix}
\]

(4.128)

Therefore, since
\[
R^{-1} = R^T
\]

(4.129)

the rotation matrix is orthogonal.

**Linear Operators:** The operator/function \( F(x) \) is a linear operator/function if
\[
F(ax) = aF(x)
\]

(4.130)

and
\[
F(x + y) = F(x) + F(y)
\]

(4.131)

If any one or these conditions is not met, then the operator/function is not linear!

**Example 11** Consider the operator that takes its argument (scalar or vector), squares it, and adds 3:
\[
O(A) = A^2 + 3
\]

(4.132)

Is this a linear operator?
Solution: We need to check the two conditions stated above. First condition:

\[ O(aA) = (aA)^2 + 3, \quad (4.133) \]

\[ aO(A) = a(A^2 + 3) = aA^2 + 3a \Rightarrow O(aA) \neq aO(A) \]

You can conclude the function is not linear because if one of the conditions is violated then the function is not linear. But let’s check also the second condition,

\[ O(A+B) = (A + B)^2 + 3 \Rightarrow O(A+B) = A^2 + AB + BA + B^2 + 3 \quad (4.134) \]

\[ O(A) + O(B) = A^2 + 3 + B^2 + 3 \Rightarrow O(A) + O(B) = A^2 + B^2 + 6 \quad (4.135) \]

which shows

\[ O(A + B) \neq O(A) + O(B) \quad (4.136) \]

### 4.9 Linear independence

A set of linear equations or vectors could be linearly dependent or independent. In order to obtain a complete solution to a set of linear equations using matrices the equations must be linearly independent. Sometimes in some physical problems it may be necessary to express a given vector as a set of linearly independent vectors. In this section we will see how linear independency in a set of linear equations, vectors, and functions can be verified and how we can obtain a set of linearly independent orthogonal vectors.

**Linear independency-set of linear equations:** The existence of the solutions to a set of linear equations with \( n \) unknowns depends on the linear independency of the equations which is determined by the Rank of the Augmented \((A)\) and Coefficient matrices \((M)\). The **Rank of a Matrix** is the number of nonzero rows remaining when a matrix has been row reduced.

- **a.** If \((\text{rank } M) < (\text{rank } A)\), the equations are inconsistent and there is no solution.
- **b.** If \((\text{rank } M) = (\text{rank } A) = n \) (number of unknowns), there is one solution.
- **c.** If \((\text{rank } M) = (\text{rank } A) = m < n\), then \( m \) unknowns can be found in terms of the remaining \( n - m \) unknowns.

**Example 12** [Linear independence-set of linear equations: in this example the current rule is excluded in order to illustrate linear independency in a set of linear equations.] Consider the circuit shown below. The circuit element values are \( R = 1.0 \Omega \) and \( \Delta V = 2.0 V \). Find the values of the three branch currents \( I_1, I_2, \) and \( I_3 \) by solving the system of three simultaneous equations obtained by applying only Kirchhoff’s voltage law to three different loops.
Solution: Applying Kirchhoff’s voltage law to three different loops we can write

\[ \Delta V + I_1 R + 2I_2 R - 2\Delta V + I_1 R = 0 \Rightarrow 2I_1 R + 2I_2 R = \Delta V, \]  
(4.137)

\[ 2\Delta V - 2I_2 R + I_3 R - 2\Delta V + I_3 R = 0 \Rightarrow -2I_2 R + 2I_3 R = 0, \]  
(4.138)

and

\[ \Delta V + I_1 R + I_3 R - 2\Delta V + I_3 R + I_1 R = 0 \Rightarrow 2I_1 R + 2I_3 R = \Delta V. \]  
(4.139)

Substituting the values for \( R \) and \( \Delta V \), we find

\[ I_1 + I_2 = 1, -I_2 + I_3 = 0, I_1 + I_3 = 1, \]  
(4.140)

which may be put in a matrix form

\[
\begin{pmatrix}
1 & 1 & 0 \\
0 & -1 & 1 \\
1 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
I_1 \\
I_2 \\
I_3
\end{pmatrix}
= 
\begin{pmatrix}
1 \\
0 \\
1
\end{pmatrix},
\]

(4.141)

There follows that for the Augmented matrix, we have

\[
A = 
\begin{pmatrix}
1 & 1 & 0 & 1 \\
0 & -1 & 1 & 0 \\
1 & 0 & 1 & 1
\end{pmatrix}
\]

(4.142)

and for the coefficient matrix

\[
M = 
\begin{pmatrix}
1 & 1 & 0 \\
0 & -1 & 1 \\
1 & 0 & 1
\end{pmatrix}
\]

(4.143)

To find the Rank of the Augmented and the Coefficient matrix we carry out elementary row operation:

Add row two to row one:

\[
\begin{pmatrix}
1 & 0 & 1 & 1 \\
0 & -1 & 1 & 0 \\
1 & 0 & 1 & 1
\end{pmatrix}
\]

(4.144)
and subtract row three from row one

\[
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & -1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
\end{pmatrix}.
\] (4.145)

We note that the Augmented matrix can not be row reduced any further. The rank of the coefficient matrix is 2 and also the rank of the augmented matrix is 2. That means \( \text{rank } M = \text{rank } A \), but the number of unknown variables is 3. This means only two unknowns can be found in terms of the third unknown.

**Linearly dependent and independent vectors:** If we define three vectors from the three equations we obtained in the example above

\[
\vec{A} = (1, 1, 0), \vec{B} = (0, -1, 0), \vec{C} = (1, 0, 1)
\] (4.146)

we can see that

\[
\vec{A} + \vec{B} - \vec{C} = 0
\] (4.147)

These three vectors are linearly dependent vectors. In general for a set of vectors \( \vec{A}_1, \vec{A}_2, \vec{A}_3, ..., \vec{A}_n \), if

\[
k_1 \vec{A}_1 + k_2 \vec{A}_2 + k_3 \vec{A}_3 + ... + k_n \vec{A}_n = 0,
\] (4.148)

where \( k_i \) is none zero real number, the vectors are said to be linearly dependent otherwise the vectors are linearly independent. Let’s consider three dimensional, three vectors

\[
\vec{A}_1 = a_1 \hat{x} + a_2 \hat{y} + a_3 \hat{z},
\]

\[
\vec{B}_1 = b_1 \hat{x} + b_2 \hat{y} + b_3 \hat{z},
\]

\[
\vec{C}_3 = c_1 \hat{x} + c_2 \hat{y} + c_3 \hat{z}.
\]

Using matrix representation, we may write

\[
k_1 \vec{A}_1 + k_2 \vec{A}_2 + k_3 \vec{A}_3 + ... + k_n \vec{A}_n = 0
\]

\[
\Rightarrow \begin{pmatrix}
k_1 a_1 & k_1 a_2 & k_1 a_3 \\
k_2 b_1 & k_2 b_2 & k_2 b_3 \\
k_3 c_1 & k_3 c_2 & k_3 c_3 \\
\end{pmatrix}
\begin{pmatrix}
\hat{x} \\
\hat{y} \\
\hat{z}
\end{pmatrix} = 0.
\] (4.150)

This equality holds if and only if

\[
\det \begin{vmatrix}
k_1 a_1 & k_1 a_2 & k_1 a_3 \\
k_2 b_1 & k_2 b_2 & k_2 b_3 \\
k_3 c_1 & k_3 c_2 & k_3 c_3 \\
\end{vmatrix} = 0
\] (4.151)

\textit{N.B.} This condition is applicable if and only if the dimension of the vector is the same as the number of the vectors. In other words when the coefficient matrix forms a square matrix.
Example 13 Consider the set of three mutually orthogonal Cartesian unit vectors. Prove that these three unit vectors form a linearly independent set of vectors.

Solution: For the three unit vectors we may write that

\[ \vec{A}_1 = \hat{x}, \Rightarrow a_1 = 1, a_2 = a_3 = 0 \]  
\[ \vec{B}_1 = \hat{y}, b_2 = 1, b_1 = b_3 = 0 \]  
\[ \vec{C}_3 = \hat{z} \Rightarrow c_1 = c_2 = 0, c_3 = 1 \]  

so that

\[
\begin{vmatrix}
    k_1a_1 & k_1a_2 & k_1a_3 \\
    k_2b_1 & k_2b_2 & k_2b_3 \\
    k_3c_1 & k_3c_2 & k_3c_3
\end{vmatrix}
= \det
\begin{vmatrix}
    k_1 & 0 & 0 \\
    0 & k_2 & 0 \\
    0 & 0 & k_3
\end{vmatrix}
= k_1 k_2 k_3.
\]  

Therefore

\[
\begin{vmatrix}
    k_1a_1 & k_1a_2 & k_1a_3 \\
    k_2b_1 & k_2b_2 & k_2b_3 \\
    k_3c_1 & k_3c_2 & k_3c_3
\end{vmatrix}
= k_1 k_2 k_3,
\]  

can be zero if and only if \( k_1 = 0, k_2 = 0, \) or \( k_3 = 0 \). This means that the Cartesian unit vectors are **linearly independent**.

In general, any set of **linearly independent vectors** either is already a **mutually orthogonal set of vectors** or else the vectors in the set can be combined to form a set of mutually orthogonal vectors (in a process called Gram-Schmidt orthogonalization).

**A Basis Set of Vectors**: these are set of linearly independent vectors which span a vector space.

**A Spanning Set of Vectors**: a set of vectors spans a space if all the vectors in the space can be written as a linear combination of the spanning **basis set of vectors**.

Example 14 Consider the four set of vectors given by

\[ \vec{A} = (1, 4, -5), \vec{B} = (5, 2, 1), \]  
\[ \vec{C} = (2, -1, 3), \vec{D} = (3, -6, 11) \]  

and the two dimensional vector space (a plane) formed by the two vectors

\[ \vec{r}_1 = (9, 0, 7), \vec{r}_2 = (0, -9, 13) \]

and the origin. Show that

(a) The four vectors \( \vec{A}, \vec{B}, \vec{C}, \) and \( \vec{D} \) are a spanning set of vectors in the two dimensional space formed by the two vectors \( \vec{r}_1 \) and \( \vec{r}_2 \).

(b) Show that the vectors \( \vec{r}_1 \) and \( \vec{r}_2 \) are a basis set of vectors for the two dimensional vector space.
Solution: If we perform row reduction for the matrix formed by the four vectors,
\[
M = \begin{bmatrix}
1 & 4 & -5 \\
5 & 2 & 1 \\
2 & -1 & 3 \\
3 & -6 & 11
\end{bmatrix}
\]
we find
\[
M' = \begin{bmatrix}
9 & 0 & 7 \\
0 & -9 & 13 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]
(4.158)
All the row operations are reversible and we can write the vectors
\[
\vec{A} = (1, 4, -5), \vec{B} = (5, 2, 1),
\]
\[
\vec{C} = (2, -1, 3), \vec{D} = (3, -6, 11)
\]
in terms of the basis set of vectors
\[
\vec{r}_1 = (9, 0, 7), \vec{r}_2 = (0, -9, 13)
\]
(4.160)

4.10 Gram-Schmidt orthogonalization

We call a set of vectors orthonormal if they are mutually orthogonal (perpendicular), and each vector is normalized (that is its norm is one). For example the vectors \(\hat{x}, \hat{y}, \hat{z}\) form an orthonormal set. Gram-Schmidt orthogonalization is a systematic process of obtaining orthonormal basis vectors from a set of basis vectors. Suppose we have basis vectors \(\vec{A}, \vec{B}, \vec{C}\), we can obtain an orthonormal set of basis vectors applying the following procedure:

1. Normalize \(\vec{A}\)- you get the first orthonormal basis vector
\[
\hat{A} = \frac{\vec{A}}{|\vec{A}|}.
\]
(4.161)

2.

(a) Find the component of the second vector, \(\vec{B}\), along the direction of \(\hat{A}\)
\[
\vec{B}_A = \left(\vec{B} \cdot \hat{A}\right) \hat{A}.
\]
(4.162)

(b) Subtract this component from \(\vec{A}\) and normalize the resulting vector,
\[
\hat{B} = \frac{\vec{B} - \left(\vec{B} \cdot \hat{A}\right) \hat{A}}{|\vec{B} - \left(\vec{B} \cdot \hat{A}\right) \hat{A}|}.
\]
(4.163)
This is the second orthonormal vector in the set. It is orthogonal to \(\hat{A}\) and it has a unit magnitude.
3. 
(a) Find the components of the third vector, $\vec{C}$, along the directions of $\vec{A}$ and $\vec{B}$

$$\vec{C}_A = (\vec{C} \cdot \hat{A}) \hat{A}, \vec{C}_B = (\vec{C} \cdot \hat{B}) \hat{B}. \quad (4.164)$$

(b) Subtract these components from $\vec{C}$ and normalize the resulting vector,

$$\hat{C} = \frac{\vec{C} - (\vec{C} \cdot \hat{A}) \hat{A} - (\vec{C} \cdot \hat{B}) \hat{B}}{|\vec{C} - (\vec{C} \cdot \hat{A}) \hat{A} - (\vec{C} \cdot \hat{B}) \hat{B}|}. \quad (4.165)$$

This is the third orthonormal vector in the set. It is orthogonal to both $\hat{A}$ and $\hat{B}$ it also has a unit magnitude.

4. For a space of higher dimensions, we continue to implement this procedure.

**Example 15** For the given set of basis vectors, use Gram-Schmidt method to find an orthonormal set

$$\vec{A} = (0, 2, 0, 0), \vec{B} = (3, -4, 0, 0), \vec{C} = (1, 2, 3, 4). \quad (4.166)$$

**Solution:** Following the first step we normalize vector $\hat{A}$

$$\hat{A} = \frac{\vec{A}}{|\vec{A}|} = (0, 1, 0, 0). \quad (4.167)$$

and using step 2(a), we have

$$\vec{B}_A = (\vec{B} \cdot \hat{A}) \hat{A} = [(3, -4, 0, 0) \cdot (0, 1, 0, 0)] (0, 1, 0, 0) = (0, -4, 0, 0) \quad (4.168)$$

so that following step 2(b) the second orthonormal vector in the set can be expressed

$$\hat{B} = \frac{\vec{B} - (\vec{B} \cdot \hat{A}) \hat{A}}{|\vec{B} - (\vec{B} \cdot \hat{A}) \hat{A}|} = \frac{(3, -4, 0, 0) - (0, -4, 0, 0)}{|(3, -4, 0, 0) - (0, -4, 0, 0)|} = \frac{(3, 0, 0, 0)}{|(3, 0, 0, 0)|} = (1, 0, 0, 0) \quad (4.169)$$

Using

$$\vec{C} \cdot \hat{B} = (1, 2, 3, 4) \cdot (1, 0, 0, 0) = 1$$
$$\vec{C} \cdot \hat{A} = (1, 2, 3, 4) \cdot (0, 1, 0, 0) = 2 \quad (4.170)$$

one finds

$$\vec{C} - (\vec{C} \cdot \hat{A}) \hat{A} - (\vec{C} \cdot \hat{B}) \hat{B} = (1, 2, 3, 4) - 2 (0, 1, 0, 0) - (1, 0, 0, 0) = (0, 0, 3, 4).$$
The third orthonormal basis vector, following step 3(a) and (b), becomes

$$\hat{C} = \frac{C - (\hat{C} \cdot \hat{A}) \hat{A} - (\hat{C} \cdot \hat{B}) \hat{B}}{|C - (\hat{C} \cdot \hat{A}) \hat{A} - (\hat{C} \cdot \hat{B}) \hat{B}|} = \left( \begin{array}{c} 0, 0, 3, 4 \end{array} \right) = \left( \begin{array}{c} 0, 3, 4 \end{array} \right)$$

(4.171)

Function Spaces and Linear Independence of Sets of Functions: If $f_1(x), f_2(x), f_3(x), \ldots f_n(x)$ have derivatives of order $n - 1$, and if the determinant

$$W = \det \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f'_1(x) & f'_2(x) & \cdots & f'_n(x) \\ \vdots & \vdots & \ddots & \vdots \\ f^{n-1}_1(x) & f^{n-1}_2(x) & \cdots & f^{n-1}_n(x) \end{vmatrix} \neq 0$$

(4.172)

then the functions are linearly independent. The determinant $W$ is called the Wronskian of the functions.

Example 17 Is the set of functions $\{\sin(\omega t), \sin^2(\omega t)\}$ a linearly independent set? ($\omega = \text{constant}$)

Solution: We can write the Wronskian for the two functions as

$$W = \det \begin{vmatrix} f_1(x) & f_2(x) \\ f'_1(x) & f'_2(x) \end{vmatrix} = \det \begin{vmatrix} \sin(\omega t) & \sin^2(\omega t) \\ \frac{d}{dt}[\sin(\omega t)] & \frac{d}{dt}[\sin^2(\omega t)] \end{vmatrix}$$

$$= \det \begin{vmatrix} \sin(\omega t) & \sin^2(\omega t) \\ \omega \cos(\omega t) & 2\omega \sin(\omega t) \cos(\omega t) \end{vmatrix}$$

$$= 2\omega \sin^2(\omega t) \cos(\omega t) - \omega \sin^2(\omega t) \cos(\omega t)$$

$$W = \omega \sin^2(\omega t) \cos(\omega t) \neq 0$$

and the functions are independent.

4.11 Homework Assignment 7

1. Solve the following set of equations for the four unknowns by reducing the matrix using elementary row operations.

$$x - 2y - 4 = 0, 2z + w = 5, w + y = 2, 3x - 1 = 5z.$$  

2. Solve the following dc circuit problems from the algebra-based physics class using row reduction of matrices

(a) The circuit diagram for this problem is shown in the figure below
3. (a) What is meant by the specific gravity (s.g) of a substance?

(b) What is Archimedes’ Principle? Provide a clear statement of this principle.

(c) An object is composed of \( x \) grams of lead (\( s_g = 11 \)) and \( y \) grams of tin (\( s_g = 7 \)). The object has a mass of 82 grams in air, and an apparent mass of 77 grams when suspended in oil having a specific gravity of 0.5. Find the values of \( x \) and \( y \).

4. The half-life of a radioactive substance is the time it takes for half of it to decay into some other substances called decay products. Suppose that a radioactive sample consists of components \( A \) and \( B \) having half-lives 2 hr and 3 hr, respectively. Assume that the decay products are gases that escape at once. At the end of 12 hr the sample is found to have a mass of 56 grams, and is found to have a mass of 12 grams at the end of 18 hr. Find the masses of \( A \) and \( B \) that were originally present in the sample.

5. Do all the problems in the sample test. No need of turning your solutions in!

4.12 Homework Assignment 8

1.
(a) Given the following sets of equations

\[ x' = \frac{1}{2} \left( x + y\sqrt{3} \right), \quad y' = \frac{1}{2} \left( -x\sqrt{3} + y \right), \quad \text{and} \quad z' = z \]

and

\[ x'' = \frac{1}{2} \left( -x' + y'\sqrt{3} \right), \quad y'' = -\frac{1}{2} \left( x'\sqrt{3} + y' \right), \quad \text{and} \quad z'' = z' \]

write each set as a matrix equation and then solve for \( x'' \) and \( y'' \) in terms of \( x \) and \( y \) by multiplying matrices. Theses equations represent rotations of axes about the z-axis. By comparing the equations and matrices with the rotation matrix derived in class, find the rotation angles and check your result.

(b) Prove that the matrices corresponding to the two sets of equations given in (a) above can correspond to rotation matrices by demonstrating that they are orthogonal matrices.

2. Show that the function

\[ f(\vec{r}) = \vec{A} \cdot \vec{r} + 3 \]

is linear or none linear function.

3. Show that the vector function

\[ f(\vec{r}) = \vec{A} \times \vec{r} \]

is a linear or none linear vector function.

4. Define the integral with respect to \( x \) from 0 to 1; the object being operated on are function of \( x \). Show that this integral operator is linear or none linear operator. (This means we want to show that the operator

\[ \int_0^1 \mathcal{F}(x) \, dx \]

is a linear or none linear operator.)

5.

(a) For the matrix

\[
A = \begin{bmatrix}
1 & 0 & 5i \\
-2i & 2 & 0 \\
1 & 1 + i & 0
\end{bmatrix}
\]

find the transpose, the inverse, the complex conjugate, and the transpose conjugate.

(b) Show that the product \( AA^T \) is a symmetric matrix.
4.13 Homework Assignment 9

1. For each of the following problems write and row reduce the augmented matrix to find out whether the given set of equations have exactly one solution, no solution, or an infinite set of solutions.

(a)  
\[
\begin{align*}
2x + y - z &= 2 \\
4x + y - 2z &= 3
\end{align*}
\]

(b)  
\[
\begin{align*}
2x + 5y + z &= 2 \\
x + y + 2z &= 1 \\
x + 5z &= 3
\end{align*}
\]

2. Find the rank of the matrix
\[
M = \begin{pmatrix}
1 & 1 & 4 & 3 \\
3 & 1 & 10 & 7 \\
4 & 2 & 14 & 10 \\
2 & 0 & 6 & 4
\end{pmatrix}
\]

3. Write the vectors
\[
\vec{A} = (1, 4, -5), \vec{B} = (5, 2, 1), \vec{C} = (2, -1, 3), \vec{D} = (3, -6, 11)
\]
as a linear combination of the vectors
\[
\vec{a} = (9, 0, 7), \vec{b} = (0, -9, 13)
\]

4. In problem (a) and (b) show that the given functions are linearly independent functions.

(a)  
\[
f_1 (x) = \sin (x), f_2 (x) = \cos (x)
\]

(b)  
\[
f_1 (x) = x, f_2 (x) = e^x, f_3 (x) = xe^x
\]

5. Show that the matrix
\[
M = \begin{pmatrix}
\frac{1+i\sqrt{3}}{4} & \frac{\sqrt{3}}{2\sqrt{2}} (1 + i) \\
-\frac{\sqrt{3}}{2\sqrt{2}} (1 + i) & \frac{\sqrt{3}+i}{4}
\end{pmatrix}
\]
is a unitary matrix. That means we need to show that
\[
M^{-1} = M^T.
\]
Chapter 5

Introduction to differential calculus-I

In advanced undergraduate or graduate physics courses we often deal with differential calculus. In this chapter we will get introduced to the basics in differential calculus. These includes partial differentiation and total differential of multivariable functions and its applications to real physical problems.

5.1 Partial differentiation

Consider a well-defined function \( z(x, y) \) that defines the \( z \) coordinate of a point in space. This function depends on two coordinates (\( x \) and \( y \)). This function is differentiable with respect to both \( x \) and \( y \) for all \( x \) and \( y \). In the case where \( z \) depends only one coordinates, for example \( z(x) \), the differentiation of this function with respect to \( x \) is given by

\[
\frac{dz}{dx} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x \to 0} \left[ \frac{z(x + \Delta x) - z(x)}{\Delta x} \right] . \quad (5.1)
\]

Thus for \( z(x, y) \) the partial differentiation with respect to \( x \) or \( y \), when \( x \) and \( y \) are independent variables, is determined by keeping \( y \) or \( x \) constant and are given by

\[
\frac{\partial z(x, y)}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta f(x, y)}{\Delta x} = \lim_{\Delta x \to 0} \left[ \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \right] \quad \text{...y is kept constant} \quad (5.2)
\]

and

\[
\frac{\partial z(x, y)}{\partial y} = \lim_{\Delta y \to 0} \frac{\Delta f(x, y)}{\Delta y} = \lim_{\Delta y \to 0} \left[ \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \right] \quad \text{...x is kept constant.} \quad (5.3)
\]
Generally for a function that depends on the independent set of variables \( \{x_1, x_2, \ldots, x_n\} \), \( f(x_1, x_2, \ldots, x_n) \), the partial differentiation with respect to one of these variables, \( x_i \) is given by

\[
\frac{\partial}{\partial x_i} f(x_1, x_2, \ldots, x_n) = \lim_{\Delta x_i \to 0} \left[ \frac{f(x_1, x_2, \ldots, x_i + \Delta x_i, \ldots, x_n) - f(x_1, x_2, \ldots, x_i, \ldots, x_n)}{\Delta x_i} \right]
\]

...all \( x \neq x_i \) are kept constant \( (5.4) \)

**Example 1** Consider the function \( z(x, y) = x^2 + y^2 \), where \( x \) and \( y \) are independent.

(a) Sketch this function to show that it defines a surface in a 3-D space described by the Cartesian coordinates \( (x, y, z) \), where the \( z \) coordinate depends on the \( x \) and \( y \) coordinates.

(b) Evaluate the following partial differentiations at the point \( (x, y) = (1, 0) \):

\[
\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x} \left[ \frac{\partial z}{\partial y} \right], \frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x} \left[ \frac{\partial z}{\partial x} \right], \frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y} \left[ \frac{\partial z}{\partial x} \right] \quad (5.5)
\]

Solution:

(a) The 3-D plot of this function obtained using Mathematica is shown below
Using the function \( z(x, y) = x^2 + y^2 \), we can evaluate the partial derivatives at point \((x, y) = (1, 0)\) as:

\[
\frac{\partial z}{\partial x} = \frac{\partial}{\partial x} (x^2 + y^2) = 2x \Rightarrow \left[ \frac{\partial z}{\partial x} \right]_{x=1, y=0} = 2, \quad (5.6a)
\]

\[
\frac{\partial z}{\partial y} = \frac{\partial}{\partial y} (x^2 + y^2) = 2y \Rightarrow \left[ \frac{\partial z}{\partial y} \right]_{x=1, y=0} = 0, \quad (5.6b)
\]

\[
\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x} \left\{ \frac{\partial}{\partial y} (x^2 + y^2) \right\} = \frac{\partial}{\partial x} (2y) = 0 \Rightarrow \left[ \frac{\partial^2 z}{\partial x \partial y} \right]_{x=1, y=0} = 0. \quad (5.6c)
\]

\[
\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x} \left\{ \frac{\partial}{\partial x} (x^2 + y^2) \right\} = \frac{\partial}{\partial x} (2x) = 2 \Rightarrow \left[ \frac{\partial^2 z}{\partial x^2} \right]_{x=1, y=0} = 2. \quad (5.6d)
\]

\[
\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y} \left\{ \frac{\partial}{\partial x} (x^2 + y^2) \right\} = \frac{\partial}{\partial y} (2x) = 0 \Rightarrow \left[ \frac{\partial^2 z}{\partial y \partial x} \right]_{x=1, y=0} = 0. \quad (5.6e)
\]

**Important notation:** In some branches of physics, such as in thermodynamics and statistical physics, a function may depend on a set of variables which are not necessarily independent as they may be related by one or more variables. However, it may physically be important to find the partial differentiation with respect to one variable by keeping one or more of the other variables constant. In such cases, the partial differentiation is expressed using subscripts for the variables that need to be kept constant. For example

\[
\left( \frac{\partial z}{\partial r} \right)_x
\]

mean "the partial of \( z \) with respect to \( r \), with \( x \) held constant" with \( z \) expressed as a function of \( r \) and \( x \).

**Example 2a** For \( z(x, y) = x^2 + 2y^2 \), \( x = r \cos(\theta) \), and \( y = r \sin(\theta) \), find the following partial derivative

\[
\left( \frac{\partial z}{\partial r} \right)_x.
\]

**Solution:** First the function \( z(x, y) \) needs to be expressed as \( z(x, r) \). To this end, we note that

\[
x^2 + y^2 = r^2 \Rightarrow y = \sqrt{r^2 - x^2}
\]

and

\[
z(x, r) = x^2 + 2y^2 = x^2 + 2(r^2 - x^2) = 2r^2 - x^2.
\]

Then one can easily show that

\[
\left( \frac{\partial z}{\partial r} \right)_x = \left( \frac{\partial}{\partial r} (2r^2 - x^2) \right)_x = 4r. \quad (5.10)
\]
Example 2b An ideal gas is defined to be a gas that obeys the ideal gas equation of state

\[ PV = nRT \] (5.11)

and

\[ U = \frac{3}{2} nRT \] (5.12)

where \( p \) is the pressure, \( V \) is the volume, \( n \) mass of the ideal gas measured in moles, \( T \) is the temperature, \( U \) is the internal energy, and \( R \) the universal gas constant. Note that \( P, V, T, \) and \( U \) are the thermodynamic variables. Obviously, as we can see from these two equations are not independent set of variables. Find expressions for the following partial differentiations:

\[ \left( \frac{\partial P}{\partial V} \right)_T, \left( \frac{\partial P}{\partial T} \right)_V, \left( \frac{\partial P}{\partial V} \right)_U, \left( \frac{\partial^2 P}{\partial V^2} \right)_T, \text{ and } \left( \frac{\partial^2 P}{\partial T^2} \right)_V. \] (5.13)

Solution: Using the ideal gas equations

\[ PV = nRT, U = \frac{3}{2} nRT \] (5.14)

we may write

\[ P = \frac{nRT}{V}, U = \frac{3}{2} nRT, T = \frac{2U}{3nR} \Rightarrow P = \frac{nR}{V} \times \frac{2U}{3nR} = \frac{2U}{3V} \] (5.15)

so that

\[ \left( \frac{\partial P}{\partial V} \right)_T = \left( \frac{\partial P(V,T)}{\partial V} \right)_T = \frac{\partial}{\partial V} \left[ \frac{nRT}{V} \right] = -\frac{nRT}{V^2}, \] (5.16a)

\[ \left( \frac{\partial P}{\partial T} \right)_V = \left( \frac{\partial P(V,T)}{\partial T} \right)_V = \frac{\partial}{\partial T} \left[ \frac{nRT}{V} \right] = \frac{nR}{V}, \] (5.16b)

\[ \left( \frac{\partial P}{\partial V} \right)_U = \left( \frac{\partial P(V,U)}{\partial V} \right)_U = \frac{\partial}{\partial V} \left[ \frac{2U}{3V} \right] = \frac{2U}{3V^2}, \] (5.16c)

\[ \left( \frac{\partial P}{\partial U} \right)_V = \left( \frac{\partial P(U,V)}{\partial U} \right)_V = \frac{\partial}{\partial U} \left[ \frac{2U}{3V} \right] = \frac{2}{3V}, \] (5.16d)

\[ \left( \frac{\partial^2 P}{\partial V^2} \right)_T = \left( \frac{\partial^2 P(V,T)}{\partial V^2} \right)_T = \frac{\partial}{\partial V} \left[ \frac{\partial P(V,T)}{\partial V} \right] \]

\[ = \frac{\partial}{\partial V} \left[ -\frac{nRT}{V^2} \right] = \frac{2nRT}{V^3}, \] (5.17)

\[ \left( \frac{\partial^2 P}{\partial T^2} \right)_V = \left( \frac{\partial^2 P(V,T)}{\partial T^2} \right)_V = \frac{\partial}{\partial T} \left[ \frac{\partial P}{\partial T} \right] = \frac{\partial}{\partial T} \left[ \frac{nR}{V} \right] = 0 \] (5.18)
5.2 Total differential

Consider a straight wire with length $L$ that lies along the positive x-axis with one end at the origin. Imagine you glued a charge that is not uniformly distributed along the wire. That means the charge $Q$ on the wire depends on the coordinate $x$, $(Q(x))$. Thus the infinitesimal charge $\Delta Q(x)$ on the wire over the length in between points $x$ and $x + \Delta x$ depends on the length $(\Delta x)$ and how the charge changes over $\Delta x$ ($\frac{\Delta Q(x)}{\Delta x}$).

$$\Delta Q(x) = \frac{\Delta Q(x)}{\Delta x} \Delta x. \quad (5.19)$$

In the limit as the interval length, $\Delta x$, goes to zero, we note

$$\frac{dQ(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta Q(x)}{\Delta x}, \quad (5.20)$$

so that the corresponding infinitesimal charge $dQ(x)$ becomes

$$dQ = \frac{dQ(x)}{dx} dx, \quad (5.21)$$

which is the differential of the function $Q(x)$. Physically, this means the infinitesimal charge on the wire in the interval $dx$ depends on how the change $Q$ changes with respect to $x$, $\left(\frac{dQ(x)}{dx}\right)$.

Consider a mountain with hills and valleys. The elevation of the mountain $z$ depends on where you are located on the mountain $(x, y)$ because of the hills and valleys. If you hike on this mountain, what you will travel on the surface of the mountain depends on what you traveled along the directions in $x$ and $y$ and also the slope of the mountain along these directions. Thus an infinitely displacement on the mountain $dz$ is given by the total differential

$$dz = \frac{\partial z(x, y)}{\partial x} dx + \frac{\partial z(x, y)}{\partial y} dy. \quad (5.22)$$

**Example 3** Application of the total differential: The height and radius of a right-circular cylinder are measured to be

$$H \pm \Delta H = 7.6 \pm 0.2\text{cm}, R \pm \Delta R = 3.7 \pm 0.2\text{cm} \quad (5.23)$$

Find an approximate value for the uncertainty in the cylinder’s volume, $dV$.

**Solution:** We recall that volume of a cylinder is given by

$$V(R, H) = \pi R^2 H \quad (5.24)$$

so that the total differential can be expressed as

$$dV = \frac{\partial V}{\partial R} dR + \frac{\partial V}{\partial H} dH \quad (5.25)$$
replacing $dV, dR, \text{ and } dH$ by $\Delta V, \Delta R, \text{ and } \Delta H$, respectively, we may express the uncertainty in volume as

$$\Delta V = \frac{\partial V}{\partial R} \Delta R + \frac{\partial V}{\partial H} \Delta H.$$  \hfill (5.26)

Evaluating

\begin{align*}
\frac{\partial V}{\partial R} &= \frac{\partial}{\partial R} [\pi R^2 H] = 2RH \quad \hfill (5.27a) \\
\frac{\partial V}{\partial H} &= \frac{\partial}{\partial H} [\pi R^2 H] = \pi R^2 \quad \hfill (5.27b)
\end{align*}

at $H = 7.6$ and $R = 3.7$, we have

\begin{align*}
\left[ \frac{\partial V}{\partial R} \right]_{H=7.6, R=3.7} &= 176.6 \quad \hfill (5.28a) \\
\left[ \frac{\partial V}{\partial H} \right]_{H=7.6, R=3.7} &= 42.9 \quad \hfill (5.28b)
\end{align*}

so that using these results along with $\Delta R = 0.2$ and $\Delta H = 0.2$, the uncertainty in the volume becomes

$$\Delta V = 176.6 \times 0.2 + 42.9 \times 0.2 = 44.76$$  \hfill (5.29)

Since uncertainty is expressed using one significant figure, we must write

$$\Delta V = 50 \quad \hfill (5.30)$$

The volume is

$$V(R, H) = \pi R^2 H = 326.7 \quad \hfill (5.31)$$

then we may write

$$V \pm \Delta V = 330 \pm 50 \text{cm}^3. \quad \hfill (5.32)$$

5.3 The multivariable form of the Chain rule

*Function of one variable*: In a function of a single variable, $y(x)$, the variable $x$ could be a function of another variable such as time, $t$, $(x(t))$. In such cases to find the time dependence of the function $y(x)$, we need to use the chain rule for a function of one variable. We recall from the previous section, the total differential for $y(x)$ can be written as

$$dy = \frac{dy}{dx} dx \quad \hfill (5.33)$$

so that diving this by $dt$, we find the Chain rule for a function of single variable

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}. \quad \hfill (5.34)$$
We often need to apply the chain rule in one, two, or three dimensions in physics.

**Function of two variables**: For a function of two variable, \( z(x, y) \), we recall that the total differential is given by

\[
dz = \frac{\partial z}{\partial x} \, dx + \frac{\partial z}{\partial y} \, dy.
\]  

(5.35)

If both \( x \) and \( y \) depend on another variable, like time, \( t \), dividing this equation by \( dt \), we find the Chain rule for a function of two variables

\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} \, \frac{dx}{dt} + \frac{\partial z}{\partial y} \, \frac{dy}{dt}.
\]  

(5.36)

This can be generalized for function of multivariable.

**Example 4** Consider the function

\[
z = \exp \left[ -x^2 - y^2 \right]
\]  

(5.37)

where

\[
x(t) = e^t \text{ and } y(t) = e^{-t}.
\]  

(5.38)

Find

\[
\frac{dz}{dt}.
\]  

(5.39)

**Solution**: Since \( x \) and \( y \) are a function of the same variable, \( t \), we can write

\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}.
\]  

(5.40)

Noting that

\[
\frac{\partial z}{\partial x} = (-2x) \exp \left[ -x^2 - y^2 \right] \\
\frac{\partial z}{\partial y} = (-2y) \exp \left[ -x^2 - y^2 \right] \\
\frac{dx}{dt} = e^t = x, \quad \frac{dy}{dt} = -e^t = -y
\]  

we find

\[
\frac{dz}{dt} = (-2x) \exp \left[ -x^2 - y^2 \right] x + (-2y) \exp \left[ -x^2 - y^2 \right] (-y)
\]  

\[
\Rightarrow \frac{dz}{dt} = 2 \left( y^2 - x^2 \right) \exp \left[ -x^2 - y^2 \right]
\]  

(5.42)

**Example 5** Show that the trajectory of the projectile motion of a ball with mass \( m \) thrown from the ground \((x_0 = 0, y_0 = 0)\) at an angle \( \theta \) (measured from the horizontal (+x-axis)) with a speed \( v_0 \) is defined by equation of hyperbola given by

\[
y(x) = bx - ax^2,
\]  

(5.43)
CHAPTER 5. INTRODUCTION TO DIFFERENTIAL CALCULUS-I

where

\[ a = \frac{g}{2v_0 \cos^2(\theta)}, b = v_0 \tan(\theta). \]  

(5.44)

Note that \( x \) is the distance of the ball measured on the ground and \( y(x) \) is the height of the ball from the ground.

**Solution:** From Newton’s second law

\[ F_x = ma_x \text{ and } F_y = ma_y \]  

(5.45)

Since for a projectile motion there is no force on the \( x \) direction, \( F_x = 0 \), we find

\[ a_x = \frac{dv_x}{dt} = 0 \Rightarrow v_x(t) = v_{0x} \Rightarrow \frac{dx}{dt} = v_{0x} \Rightarrow x(t) = x_0 + v_{0x}t \]

(5.46)

This shows that the \( x \) coordinate depends on time \( t \). In the \( y \)–direction there is gravitational force, \( F_y = -mg \) that leads to

\[ a_y = \frac{dv_y}{dt} = -g \Rightarrow v_y(t) = v_{0y} - gt \Rightarrow \frac{dy}{dt} = v_0 \sin(\theta) - gt. \]  

(5.47)

Using the chain rule for a function of one variable, one can write

\[ \frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}. \]  

(5.48)

From the result in Eq. (5.46), we have

\[ \frac{dx}{dt} = v_0 \cos(\theta) \text{ and } t = \frac{x(t) - x_0}{v_0 \cos(\theta)} \]  

(5.49)

so that substituting these equations along with the result in Eq. (5.47), one may write Eq. (5.48) as

\[ v_0 \sin(\theta) - g \left( \frac{x(t) - x_0}{v_0 \cos(\theta)} \right) = \frac{dy}{dx} v_0 \cos(\theta) \]  

(5.50)

There follows that

\[ \frac{dy}{dx} = v_0 \tan(\theta) - \frac{g}{v_0 \cos^2(\theta)} (x - x_0). \]  

(5.51)

Upon integrating this equation with respect to \( x \),

\[ \int dy = \int \left[ v_0 \tan(\theta) - \frac{g}{v_0 \cos^2(\theta)} (x - x_0) \right] dx \]  

(5.52)

we find

\[ y(x) = y_0 + v_0 \tan(\theta) x - \frac{g}{v_0 \cos^2(\theta)} \left( \frac{x^2}{2} - x_0 x \right). \]  

(5.53)
5.3. THE MULTIVARIABLE FORM OF THE CHAIN RULE

Assuming the initial position of the projectile is at the origin \((x_0 = 0, y_0 = 0)\), this can be put in the form

\[
y(x) = bx - ax^2,
\]

where

\[
a = \frac{g}{2v_0 \cos^2 \theta}, \quad b = v_0 \tan \theta.
\]

**Example 6** Implicit Differentiation and application of the total differential:
Consider the equation

\[
y^3 - x^2y = 8
\]

where \(y\) is a function of \(x\). Evaluate \(dy/dx\) and \(d^2y/dx^2\) at the point \((x, y) = (3, -1)\), where \(y = y(x, y)\)

**Solution:** The easy way to find both

\[
\frac{dy}{dx} \quad \text{and} \quad \frac{d^2y}{dx^2}
\]

is to apply implicit differentiation. To this end, let

\[
z(x, y) = y^3 - x^2y = 8
\]

so that

\[
\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \frac{dy}{dx} = 0 \Rightarrow \frac{dz}{dx} = -2xy + (3y^2 - x^2) \frac{dy}{dx} = 0
\]

\[
\Rightarrow \frac{dy}{dx} = -\frac{2xy}{3y^2 - x^2}
\]

Then at \((x, y) = (3, -1)\), we find

\[
\frac{dy}{dx} = 1.
\]

To find the second derivative of the function we use the result for the first derivative. Noting that

\[
\frac{dy}{dx} = \frac{2xy}{3y^2 - x^2} = f(x, y)
\]

we can write

\[
\frac{d^2y}{dx^2} = \frac{\partial}{\partial x} f(x, y) + \left[ \frac{\partial}{\partial y} f(x, y) \right] \frac{dy}{dx}.
\]

Upon carrying out the partial differentiations, we have

\[
\frac{\partial}{\partial x} [f(x, y)] = \frac{2y}{3y^2 - x^2} + \frac{4x^2 y}{(3y^2 - x^2)^2}
\]
and 
\[ \frac{\partial}{\partial y} [f(x, y)] = \frac{2x}{3y^2 - x^2} - \frac{12xy^2}{(3y^2 - x^2)^2}, \] (5.64)
so that Eq. (5.62) becomes
\[ \frac{d^2y}{dx^2} = \frac{2y}{3y^2 - x^2} + \frac{4x^2y}{(3y^2 - x^2)^2} + \left( \frac{2x}{3y^2 - x^2} - \frac{12xy^2}{(3y^2 - x^2)^2} \right) \frac{dy}{dx}. \] (5.65)
This can be rewritten as
\[ \frac{d^2y}{dx^2} = \frac{1}{x} \left( \frac{2xy}{3y^2 - x^2} \right) + \frac{1}{y} \left( \frac{2xy}{3y^2 - x^2} \right)^2 \]
\[ + \left[ \frac{1}{y} \left( \frac{2xy}{3y^2 - x^2} \right) - \frac{3}{x} \left( \frac{2xy}{3y^2 - x^2} \right)^2 \right] \frac{dy}{dx} \] (5.66)
and using the result in Eq. (5.59), we find
\[ \frac{d^2y}{dx^2} = \frac{1}{x} \frac{dy}{dx} + \frac{1}{y} \left( \frac{dy}{dx} \right)^2 + \left[ \frac{1}{y} \left( \frac{dy}{dx} \right) - \frac{3}{x} \left( \frac{dy}{dx} \right)^2 \right] \frac{2xy}{3y^2 - x^2} \] (5.67)
Recalling the result we obtain for Eq. (5.59) at \((x, y) = (3, -1)\)
\[ \frac{dy}{dx} = 1 \] (5.68)
we find from Eq. (5.122)
\[ \frac{d^2y}{dx^2} = \frac{1}{3} - 1 - 1 = \frac{-8}{3}. \] (5.69)

### 5.4 Extremum (Max/Min) problems

For a function of single variable \(y(x)\), the first derivative
\[ f(x) = \frac{dy}{dx} \] (5.70)
tells us the slope of the tangent line at the point on the curve with the coordinate \((x, y(x))\). The tangent line to an extremum point on a curve is a horizontal line, the slope of which is zero. Therefore, if the \(x\) coordinates of the extremum point on the curve defined by \(y(x)\) is \(x_i\), since the slope of the tangent line at \((x_i, y(x_i))\) is zero,
\[ f(x_i) = \frac{dy}{dx} \bigg|_{x=x_i} = 0. \] (5.71)

For a function of two variable, \(z(x, y)\), the extremum points with coordinates \((x_i, y_i, z(x_i, y_i))\) represent the coordinates for maxima on the hills or the minima
5.4. EXTREMUM (MAX/MIN) PROBLEMS

in the valleys on the surface defined by the function $z(x, y)$. The tangent plane to these extremum points are horizontal planes (parallel to the x-y plane). For this plane the slopes along the $x$ and the $y$ must be zero. Mathematically, these slopes are defined by the partial derivative of the function $z(x, y)$ evaluated at the maxima or minima points with coordinates $(x_i, y_i)$. Thus at the extremum (maxima or minima) points

$$\frac{\partial f(x, y)}{\partial x} \bigg|_{x=x_i} = 0 \quad \text{and} \quad \frac{\partial f(x, y)}{\partial y} \bigg|_{y=y_i} = 0.$$ (5.72)

These two equations defines the horizontal tangent plane at the extremum points on the surface defined by $z(x, y)$.

**Example 7** An aquarium with rectangular sides and bottom has a fixed volume $V_0$. Find its proportions so that it will require the least amount of glass for construction.

**Solution:** As shown in Fig. 5.1, the length, width, and height of the aquarium are $l, w, \text{and } h$, respectively. The the volume, which is fixed, is related to $l, w, \text{and } h$ by

$$V(l, w, h) = lwh = V_0.$$ (5.73)

The aquarium has rectangular sides made of glass. The surface area, $A(l, w, h)$ of the five sides (excluding the top side which is open) is given by

$$A(l, w, h) = lw + 2(lh + wh).$$ (5.74)

We are interested in constructing an aquarium with least amount of glass. It means the area of the glass sides that the aquarium is constructed from must be minimized while the volume remains constant. As we can see
from Eq. (5.1) \( l, w, \) and \( h \) are not independent variables. We can reduce the area from a function of three variable to a three variable by expressing \( h \) in terms of \( l \) and \( w \) using Eq. (5.1) as

\[
h = \frac{V_0}{lw}.
\]  
(5.75)

Substituting Eq. (5.75) into Eq. (5.74) we can then express the area in terms of two independent variables

\[
A(l, w) = lw + 2V_0 \left( \frac{1}{l} + \frac{1}{w} \right).
\]  
(5.76)

For the minimum area we must have

\[
\frac{\partial A(l, w)}{\partial l} = 0
\]  
(5.77)

and

\[
\frac{\partial A(l, w)}{\partial w} = 0
\]  
(5.78)

so that using Eq. (5.77), one finds

\[
w - \frac{2V_0}{l^2} = 0
\]  
(5.79)

and

\[
l - \frac{2V_0}{w^2} = 0
\]  
(5.80)

respectively. Now combining these two equations, we find

\[
w - \frac{2V_0}{w^4} = 0 \Rightarrow w \left( 1 - \frac{w^3}{2V_0} \right) = 0 \Rightarrow w = 0, w = (2V_0)^{1/3}
\]  
(5.81)

But the acceptable solution is

\[
w = (2V_0)^{1/3}.
\]  
(5.82)

Substituting combining this result and Eqs. (5.77) and (5.1) the corresponding length and height of the aquarium are

\[
l = \frac{2V_0}{w^2} = \frac{2V_0}{(2V_0)^{2/3}} = (2V_0)^{1/3}
\]  
(5.83)

\[
h = \frac{V_0}{(2V_0)^{1/3} \cdot (2V_0)^{1/3}} = \left( \frac{V_0}{4} \right)^{1/3}
\]  
(5.84)

Thus the proportions can be expressed as

\[
\frac{l : w : h}{(2V_0)^{1/3}} = 1 : 1 : \frac{1}{8}
\]

These are the dimension of the Aquarium for least amount of glass for construction and a volume \( V_0 \).
### 5.5 The Method of Lagrangian Multipliers

In the Example 7 the volume of the rectangular aquarium, which is a function of three variables, is specified,

\[ V(l, w, h) = lwh = V_0. \]  

(5.85)

That means the volume of the rectangular aquarium given can not be less or greater than \( V_0 \). We were asked to determine the minimum total surface area,

\[ A(l, w, h) = lw + 2(lh + wh). \]  

(5.86)

This area appear to be a function three variables. But because of the condition set for the volume, we have seen that it is actually a function of two variables. Such kind of conditions are referred as constraints. The function that set the volume to \( V_0 \) in Eq. (5.85) is the constraint for the area defined by the function in Eq. (5.85).

Suppose the function \( f(x, y, z, ...) \) describes some physical observable for some system. In real physical problems the variables \( x, y, z, ... \) are subject to constraints so that they are no longer independent. It is possible, at least in principle, to use each constraint to eliminate one variable and to proceed with a new and smaller set of independent variables. The use of Lagrangian multipliers is an alternate technique that may be applied to determine extremum points for the function when this elimination of variables is inconvenient or undesirable.

Suppose we are interested in finding the extremum points for the function \( f(x, y, z) \) under the constraint defined by the function \( \phi(x, y, z) = k \), where \( k \) is a constant. We note that for the function to be extremum (including the saddle point)

\[ df = 0 \Rightarrow \frac{\partial f(x, y, z)}{\partial x} dx + \frac{\partial f(x, y, z)}{\partial y} dy + \frac{\partial f(x, y, z)}{\partial z} dz = 0 \]  

(5.87)

and the necessary and sufficient condition is

\[ \frac{\partial f(x, y, z)}{\partial x} = \frac{\partial f(x, y, z)}{\partial y} = \frac{\partial f(x, y, z)}{\partial z} = 0. \]  

(5.88)

Noting that for the constraint

\[ d\phi(x, y, z) = dk = 0 \Rightarrow \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy + \frac{\partial \phi}{\partial z} dz = 0 \]  

(5.89)

and for a function \( F(x, y, z, ...) = f(x, y, z) + \lambda \phi(x, y, z) \), for \( \lambda \) a constant, at the extremum points we may write

\[ F(x, y, z) = f(x, y, z) + \lambda \phi(x, y, z) = f(x, y, z) + \lambda k \Rightarrow dF = df = 0 \]

\[ \Rightarrow \left( \frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} \right) dx + \left( \frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} \right) dy + \left( \frac{\partial f}{\partial z} + \lambda \frac{\partial \phi}{\partial z} \right) dz = 0 \]
There follows that
\[ \frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} = 0, \quad \frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} = 0, \quad \text{and} \quad \frac{\partial f}{\partial z} + \lambda \frac{\partial \phi}{\partial z} = 0. \] (5.90)

Here we want \( x, y, \) and \( z \) so \( \lambda \) need not be determined. This method would fail if the coefficient of \( \lambda \) vanish at the extremum,
\[ \frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial z} = 0. \]

The following steps can be used to find extremum points for a function \( f(x, y, z,...) \) with a constraint \( \phi(x, y, z,...) = k \)

1. Construct the function \( F(x, y, z,...) = f(x, y, z,...) + \lambda \phi(x, y, z,...) \), where \( \lambda \) is the (unknown) Lagrangian multiplier.
2. Compute the partial derivatives of \( F \) with respect to each of the variables \( x, y, z,... \) and set them equal to zero.
3. Solve the resulting system of simultaneous equations generated in Step 2 along with the constraint equation
\[ \phi(x, y, z,...) = constant \]

Example 8 Consider a wire bent to fit a curve defined by \( y = 1 - x^2 \) as shown in the figure below.

Suppose a string is stretched from the origin to a point \( (x, y) \) on the curve. Find the minimum length of the string.

**Solution:** We need to find the minimum value for the distance
\[ d(x, y) = \sqrt{x^2 + y^2}. \] (5.91)

with the constraint
\[ y(x) = 1 - x^2 \Rightarrow \phi(x, y) = y + x^2 = 1. \] (5.92)

Following the procedure for the Method of Lagrangian multipliers, we have
5.5. **THE METHOD OF LAGRANGIAN MULTIPLIERS**

1. Construct the function $F(x, y) = f(x, y) + \lambda \phi(x, y) = \sqrt{x^2 + y^2} + \lambda (y + x^2)$.

2. Compute the partial derivatives of $F$ with respect to each of the variables $x, y$ and set them equal to zero,
   \[
   \frac{\partial}{\partial x} F(x, y) = 0 \Rightarrow \frac{x}{\sqrt{x^2 + y^2}} + 2\lambda x = 0 \Rightarrow \frac{1}{\sqrt{x^2 + y^2}} = -2\lambda \quad (5.93)
   \]
   \[
   \frac{\partial}{\partial y} F(x, y) = 0 \Rightarrow \frac{y}{\sqrt{x^2 + y^2}} + \lambda = 0 \Rightarrow \frac{y}{\sqrt{x^2 + y^2}} = -\lambda \quad (5.94)
   \]

3. Solve the resulting system of simultaneous equations generated in Step 2 along with the constraint equation
   \[
   \phi(x, y) = \text{constant}
   \]
   to determine the coordinates $(x, y)$ at which the function $f(x, y)$ takes a maxima or minima values.
   \[
   \frac{1}{\sqrt{x^2 + y^2}} = -2\lambda , \quad \frac{y}{\sqrt{x^2 + y^2}} = -\lambda , y + x^2 = 1 \quad (5.95)
   \]

   These three equations can be combined to find the values for $x$ and $y$
   \[
   \frac{1}{\sqrt{x^2 + y^2}} = \frac{2y}{\sqrt{x^2 + y^2}}, y = 1 - x^2 \Rightarrow 1 = 2 (1 - x^2) \quad (5.96)
   \]
   \[
   \Rightarrow x = \pm \frac{1}{\sqrt{2}}, y = \frac{1}{2} \quad (5.97)
   \]

   Then the minimum distance is found to be
   \[
   d_{\text{min}} = \sqrt{\frac{3}{2}}. \quad (5.98)
   \]

**Example 8 (a) The aquarium problem**: An aquarium with rectangular sides and bottom has a fixed volume $V_0$. Find its proportions so that it will require the least amount of glass for construction.

**Solution**: In order to determine the minimum area of the glasses used using the method of Lagrangian multiplies we need to write the constraint. The constraint is the volume
   \[
   V_0 = lwh \Rightarrow \phi(l, w, h) = lwh = V_0 = \text{constant}. \quad (5.99)
   \]

The surface area of the aquarium is
   \[
   A(l, w, h) = lw + 2(wh + lh) \Rightarrow f(l, w, h) = lw + 2(wh + lh). \quad (5.100)
   \]

Constructing the function $F(l, w, h)$,
   \[
   F(l, w, h) = f(l, w, h) + \lambda \phi(l, w, h) \Rightarrow F(l, w, h) = lw + 2(wh + lh) + \lambda lwh. \quad (5.101)
   \]
We then minimize this function by partially differentiating it with respect to \( l, w, \) and \( h \)

\[
\frac{\partial F(l, w, h)}{\partial l} = 0 \Rightarrow w + 2h + \lambda wh = 0, \tag{5.102}
\]

\[
\frac{\partial F(l, w, h)}{\partial w} = 0 \Rightarrow l + 2h + \lambda lh = 0, \tag{5.103}
\]

and

\[
\frac{\partial F(l, w, h)}{\partial h} = 0 \Rightarrow 2w + 2l + \lambda lw = 0. \tag{5.104}
\]

We solve the above three equations along with the constraint

\[
lwh = V_0. \tag{5.105}
\]

to determine \( l, w, \) and \( h \). Since we do not need the solution for \( \lambda \), from Eq. (5.102) we can express \( \lambda \) as

\[
\lambda = -\frac{w + 2h}{wh} \tag{5.106}
\]

and substitute it into Eqs. (5.103) and (5.104) we find

\[
l + 2h - \left( \frac{w + 2h}{w} \right) l = 0 \Rightarrow l \left[ \frac{w + 2h}{w} - 1 \right] = 2h \Rightarrow w = l \tag{5.107}
\]

and

\[
2w + 2l - \left( \frac{w + 2h}{wh} \right) lw = 0 \Rightarrow l \left[ \frac{w + 2h}{h} - 2 \right] = 2w \Rightarrow l = 2h \tag{5.108}
\]

Now substituting these results in Eq. (5.105), we find

\[
4h^3 = V_0 \Rightarrow h = \left( \frac{V_0}{4} \right)^{1/3} \Rightarrow w = 2h = (2V_0)^{1/3} \tag{5.109}
\]

**Example 8 (b) Particle in a box—Quantum mechanics:** Consider the quantum mechanical problem of a particle with mass \( m \) in a box. The box is a rectangular box with length \( l \), width \( w \), and height \( h \). The ground state energy is given by

\[
E(l, w, h) = \frac{2\pi \hbar}{8m} \left( \frac{1}{l^2} + \frac{1}{w^2} + \frac{1}{h^2} \right).
\]

Find the proportions to sides of the box that minimize the ground state energy subject to the constraint that the volume is a constant

\[
V(l, w, h) = lwh = V_0
\]
5.6. CHANGE OF VARIABLES

Solution: Constructing the function $F(l, w, h)$,

$$F(l, w, h) = f(l, w, h) + \lambda \phi(l, w, h) \Rightarrow F(l, w, h) = \frac{\pi h}{4m} \left( \frac{1}{l^2} + \frac{1}{w^2} + \frac{1}{h^2} \right) + \lambda wh.$$  (5.110)

We then minimize this function by partially differentiating it with respect to $l$, $w$, and $h$

$$\frac{\partial F(l, w, h)}{\partial l} = 0 \Rightarrow -\frac{\pi h}{2ml^3} + \lambda wh = 0, \quad (5.111)$$

$$\frac{\partial F(l, w, h)}{\partial w} = 0 \Rightarrow -\frac{\pi h}{2mw^3} + \lambda lh = 0, \quad (5.112)$$

and

$$\frac{\partial F(l, w, h)}{\partial h} = 0 \Rightarrow -\frac{\pi h}{2mh^3} + \lambda lw = 0. \quad (5.113)$$

We solve the above three equations along with the constraint

$$lwh = V_0.$$  (5.114)

to determine $l$, $w$, and $h$. Since we do not need the solution for $\lambda$, using the first equation, we have

$$\lambda = \frac{\pi h}{2ml^3wh}.$$  (5.115)

so that using this expression for $\lambda$, one can write

$$-\frac{\pi h}{2mw^3} + \lambda lh = 0 \Rightarrow \frac{\pi h}{2mw^3} = \frac{\pi h}{2ml^2w} \Rightarrow w = l \quad (5.116)$$

and

$$-\frac{\pi h}{2mh^3} + \lambda lw = 0 \Rightarrow \frac{\pi h}{2mh^3} = \frac{\pi h}{2ml^2h} \Rightarrow h = l \quad (5.117)$$

Now substituting these results into the constraint for the volume, we find

$$l^3 = V_0 \Rightarrow l = (V_0)^{1/3} \Rightarrow w = l = h = (V_0)^{1/3} \quad (5.118)$$

The box must be a cube, $w : l : h = 1$ in order to minimize the energy of the particle.

5.6 Change of Variables

Physical problems in most branches of physics requires solving differential equations. These differential equations could be a function of single or multivariable. In order to find the solutions to these differential equations change of variables sometimes is necessary for different reasons. In order to change the differential equation from one variable to another when the differential equation is a function of multivariable, we need to apply partial differentiation of the function. In the next examples we illustrates application of partial differentiation in change of variables in a second order differential equation for single and multivariable function.
Example 9 The Bessel differential equation is given by:

\[ x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} - (1 - x) y = 0 \]  
(5.119)

In this equation \( y = y(x) \). Find the corresponding equation with the following change of variable:

\[ u = 2\sqrt{x}, y = y(u) \]  
(5.120)

Solution: Using partial differentiation, we can express

\[ \frac{dy}{dx} = \frac{\partial y}{\partial u} \frac{du}{dx} \]  
(5.121)

since \( y = y(u) \) that depends only in one variable \( u \), we may rewrite

\[ \frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} = \frac{dy}{du} \frac{1}{\sqrt{x}}. \]  
(5.122)

Using the result above the second derivative in the bessel equation can be expressed as

\[ \frac{d^2 y}{dx^2} = \frac{d}{dx} \left( \frac{dy}{du} \right) \frac{1}{\sqrt{x}} + \frac{dy}{du} \frac{d}{dx} \frac{1}{\sqrt{x}}. \]  
(5.123)

Noting that

\[ \frac{d}{dx} \left[ \frac{dy}{du} \right] = \frac{d}{du} \left[ \frac{dy}{du} \right] \frac{du}{dx} \Rightarrow \frac{d}{dx} \left[ \frac{dy}{du} \right] = \frac{d^2 y}{du^2} \frac{1}{\sqrt{x}} \]  
(5.124)

and

\[ \frac{d}{dx} \left[ \frac{1}{\sqrt{x}} \right] = -\frac{1}{2x\sqrt{x}} \]  
(5.125)

we find

\[ \frac{d^2 y}{dx^2} = \frac{1}{\sqrt{x}} \frac{d}{dx} \left[ \frac{dy}{du} \right] + \frac{dy}{du} \frac{d}{dx} \left[ \frac{1}{\sqrt{x}} \right] = \frac{1}{x} \frac{d^2 y}{du^2} - \frac{1}{2x\sqrt{x}} \frac{dy}{du}. \]  
(5.126)

Then substituting Eqs. (5.122) and (5.126) into Eq. (5.119) the Bessel equation takes the form

\[ x^2 \left( \frac{1}{x} \frac{d^2 y}{du^2} - \frac{1}{2x\sqrt{x}} \frac{dy}{du} \right) + x \left( \frac{dy}{du} \frac{1}{\sqrt{x}} \right) - (1 - x) y = 0 \]  
(5.127)

which can simplified into

\[ x \frac{d^2 y}{du^2} + \sqrt{x} \frac{dy}{du} - (1 - x) y = 0 \]  
(5.128)

From Eq. (5.120) we have \( x = u^2/4 \) and the Bessel equation becomes

\[ \frac{u^2}{4} \frac{d^2 y}{du^2} + \frac{u}{4} \frac{dy}{du} - \left( 1 - \frac{u^2}{4} \right) y = 0 \]

\[ \Rightarrow u^2 \frac{d^2 y}{du^2} + u \frac{dy}{du} + (u^2 - 4) y = 0. \]  
(5.129)
In classical or quantum mechanical studies of the electrical or magnetic properties of particles or systems in physics often involve solving a second order partial differential equation for some form of multivariable a scalar or vector function. One of these that commonly used is the Laplace equation. The Laplace equation in Cartesian coordinates for some scalar function \( V(x, y, z) \) is given by

\[
\frac{\partial^2 V(x, y, z)}{\partial x^2} + \frac{\partial^2 V(x, y, z)}{\partial y^2} + \frac{\partial^2 V(x, y, z)}{\partial z^2} = 0.
\]

In order to determine the solution to such kind of second order partial differential equations it may be necessary to change the variables from Cartesian to cylindrical \((r, \varphi, z)\) or to spherical \((r, \theta, \varphi)\) coordinates if the system has a cylindrical or spherical symmetry, respectively. Such change of variables for the given partial differential equation are determined by the equations that relates a point in space defined by the Cartesian coordinates to Cylindrical coordinates

\[
x = r \cos \theta, \quad y = r \sin (\theta), \quad z = z
\]

and to spherical coordinates

\[
x = r \cos \varphi \sin (\theta), \quad y = r \sin \varphi \sin (\theta), \quad z = r \cos \theta.
\]

In the next example we will see how we make change of variables for the two dimensional Laplace equation in Cartesian to Cylindrical coordinates.

**Example 10** In classical electrodynamics in a two-dimensional space where there is no free charges the electrical potential, \( V(x, y) \), satisfies the two-dimensional Laplace Equation given by

\[
\frac{\partial^2 V(x, y)}{\partial x^2} + \frac{\partial^2 V(x, y)}{\partial y^2} = 0
\]

Suppose the two dimensional space is a dielectric disk of radius \( R \) and you are given a boundary condition for the potential at the edge of the disk is zero \( V(x, y) = 0 \), for \( x^2 + y^2 = R^2 \). For such kind of problem it is much easier if we solve the Laplace equation in polar coordinates. Transform the Laplace equation into polar coordinates using

\[
x = r \cos (\theta), \quad y = r \sin (\theta), \quad V = V(r, \theta).
\]

**Solution:** There are two different way of transforming this equation. I will use the short way. We note that

\[
\frac{\partial V(r, \theta)}{\partial r} = \frac{\partial V(x, y)}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial V(x, y)}{\partial y} \frac{\partial y}{\partial r}
\]

so that using expressions for \( x \) and \( y \) in Eq. (5.133), one finds

\[
\frac{\partial V(r, \theta)}{\partial r} = \cos (\theta) \frac{\partial V(x, y)}{\partial x} + \sin (\theta) \frac{\partial V(x, y)}{\partial y}.
\]
Similarly, for partial differentiation with respect to \( \theta \), we have

\[
\frac{\partial V (r, \theta)}{\partial \theta} = \frac{\partial V (x, y)}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial V (x, y)}{\partial y} \frac{\partial y}{\partial \theta} = -r \sin (\theta) \frac{\partial V (x, y)}{\partial x} + r \cos (\theta) \frac{\partial V (x, y)}{\partial y}
\]  

(5.136)

dividing this equation by \( r \), one finds

\[
\frac{1}{r} \frac{\partial V (r, \theta)}{\partial \theta} = -\sin (\theta) \frac{\partial V (x, y)}{\partial x} + \cos (\theta) \frac{\partial V (x, y)}{\partial y}.
\]  

(5.137)

Now multiplying Eq. (5.135) by \( \cos (\theta) \) and Eq. (5.137) by \( \sin (\theta) \), we find

\[
\cos (\theta) \frac{\partial V (r, \theta)}{\partial r} = \cos^2 (\theta) \frac{\partial V (x, y)}{\partial x} + \sin (\theta) \cos (\theta) \frac{\partial V (x, y)}{\partial y}
\]  

(5.138)

and

\[
\frac{\sin (\theta)}{r} \frac{\partial V (r, \theta)}{\partial \theta} = -\sin^2 (\theta) \frac{\partial V (x, y)}{\partial x} + \sin (\theta) \cos (\theta) \frac{\partial V (x, y)}{\partial y}.
\]  

(5.139)

Subtracting Eq. (5.139) from Eq. (5.138) one finds

\[
\frac{\partial V (x, y)}{\partial x} = \cos (\theta) \frac{\partial V (r, \theta)}{\partial r} - \frac{\sin (\theta)}{r} \frac{\partial V (r, \theta)}{\partial \theta}.
\]  

(5.140)

Similarly, multiplying Eq. (5.135) by \( \sin (\theta) \) and Eq. (5.137) by \( \cos (\theta) \), we have

\[
\sin (\theta) \frac{\partial V (r, \theta)}{\partial r} = \sin (\theta) \cos (\theta) \frac{\partial V (x, y)}{\partial x} + \sin^2 (\theta) \frac{\partial V (x, y)}{\partial y}
\]  

(5.141)

and

\[
\frac{\cos (\theta)}{r} \frac{\partial V (r, \theta)}{\partial \theta} = -\sin (\theta) \cos (\theta) \frac{\partial V (x, y)}{\partial x} + \cos^2 (\theta) \frac{\partial V (x, y)}{\partial y}.
\]  

(5.142)

and adding these two equations, we find

\[
\frac{\partial V (x, y)}{\partial y} = \sin (\theta) \frac{\partial V (r, \theta)}{\partial r} + \frac{\cos (\theta)}{r} \frac{\partial V (r, \theta)}{\partial \theta}.
\]  

(5.143)

From Eqs. (5.140) and Eq. (5.143) follows that

\[
\frac{\partial}{\partial y} = \sin (\theta) \frac{\partial}{\partial r} + \frac{\cos (\theta)}{r} \frac{\partial}{\partial \theta},
\]  

(5.144)

\[
\frac{\partial}{\partial x} = \cos (\theta) \frac{\partial}{\partial r} - \frac{\sin (\theta)}{r} \frac{\partial}{\partial \theta}.
\]  

(5.145)
so that one can write

$$\frac{\partial^2}{\partial y^2} = \left[ \sin(\theta) \frac{\partial}{\partial r} + \frac{\cos(\theta)}{r} \frac{\partial}{\partial \theta} \right] \left[ \sin(\theta) \frac{\partial}{\partial r} + \frac{\cos(\theta)}{r} \frac{\partial}{\partial \theta} \right]$$

(5.146)

and

$$\frac{\partial^2}{\partial x^2} = \left[ \cos(\theta) \frac{\partial}{\partial r} - \frac{\sin(\theta)}{r} \frac{\partial}{\partial \theta} \right] \left[ \cos(\theta) \frac{\partial}{\partial r} - \frac{\sin(\theta)}{r} \frac{\partial}{\partial \theta} \right].$$

(5.147)

On carrying out the partial differentiation in the first square bracket on the second square bracket in Eq. (5.146) we have

$$\frac{\partial^2}{\partial y^2} = \sin^2(\theta) \frac{\partial^2}{\partial r^2} + \sin(\theta) \cos(\theta) \frac{\partial}{\partial r} \left[ \frac{1}{r} \frac{\partial}{\partial \theta} \right] + \frac{\cos(\theta)}{r^2} \frac{\partial}{\partial \theta} \left[ \sin(\theta) \frac{\partial}{\partial r} \right]$$

$$+ \frac{\cos^2(\theta)}{r^2} \frac{\partial}{\partial r} + \frac{\sin(\theta)}{r} \cos(\theta) \frac{\partial^2}{\partial \theta \partial r} - \frac{\sin(\theta) \cos(\theta)}{r^2} \frac{\partial}{\partial \theta} + \frac{\cos^2(\theta)}{r^2} \frac{\partial^2}{\partial \theta^2}.$$ 

(5.148)

so that upon performing the differentiations in the square brackets, one finds

$$\frac{\partial^2}{\partial y^2} = \sin^2(\theta) \frac{\partial^2}{\partial r^2} + \frac{2 \sin(\theta) \cos(\theta)}{r} \frac{\partial^2}{\partial r \partial \theta} - \frac{2 \sin(\theta) \cos(\theta)}{r^2} \frac{\partial}{\partial \theta}$$

$$+ \frac{\cos^2(\theta)}{r} \frac{\partial}{\partial r} + \frac{\cos^2(\theta)}{r^2} \frac{\partial^2}{\partial \theta^2}.$$ 

(5.149)

that simplifies into

$$\frac{\partial^2}{\partial y^2} = \sin^2(\theta) \frac{\partial^2}{\partial r^2} + \frac{2 \sin(\theta) \cos(\theta)}{r} \frac{\partial^2}{\partial r \partial \theta} - \frac{2 \sin(\theta) \cos(\theta)}{r^2} \frac{\partial}{\partial \theta}$$

$$+ \frac{\cos^2(\theta)}{r} \frac{\partial}{\partial r} + \frac{\cos^2(\theta)}{r^2} \frac{\partial^2}{\partial \theta^2}.$$ 

(5.150)

Following a similar procedure in Eq. (5.147), we have

$$\frac{\partial^2}{\partial x^2} = \cos^2(\theta) \frac{\partial^2}{\partial r^2} - \sin(\theta) \cos(\theta) \frac{\partial}{\partial r} \left[ \frac{1}{r} \frac{\partial}{\partial \theta} \right]$$

$$- \frac{\sin(\theta)}{r} \frac{\partial}{\partial \theta} \left[ \cos(\theta) \frac{\partial}{\partial r} \right] + \frac{\sin(\theta)}{r^2} \frac{\partial}{\partial \theta} \left[ \sin(\theta) \frac{\partial}{\partial r} \right].$$

(5.151)

that leads to

$$\frac{\partial^2}{\partial x^2} = \cos^2(\theta) \frac{\partial^2}{\partial r^2} + \frac{2 \sin(\theta) \cos(\theta)}{r} \frac{\partial^2}{\partial r \partial \theta} - \frac{2 \sin(\theta) \cos(\theta)}{r^2} \frac{\partial}{\partial \theta}$$

$$+ \frac{\sin^2(\theta)}{r} \frac{\partial}{\partial r} + \frac{\sin^2(\theta)}{r^2} \frac{\partial^2}{\partial \theta^2}.$$ 

(5.152)
Now adding Eqs. (5.150) and (5.152), we find

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}$$

which can be rewritten as

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}$$

Therefore the Laplace equation in polar coordinates can be expressed as

$$\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial V(r, \theta)}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 V(r, \theta)}{\partial \theta^2} = 0$$

5.7 Legendre Transformations

*Legendre Transformation* of a function $f(x)$ denoted by $g(p)$ is given by

$$g(p) = xp - f(x)$$

where

$$p = \frac{df}{dx}$$

The Legendre transformation in Eq. (5.156) can also be expressed using Eq. (5.157) as

$$g \left( \frac{df}{dx} \right) = x \frac{df}{dx} - f(x)$$

It is commonly used in thermodynamics and in the Hamiltonian formulation of classical mechanics. Here we will consider an example in thermodynamics. As mentioned previously, in thermodynamics we often have many variables to choose from in order to specify the state of a given system, but only a small number of these variables are actually independent variables. So-called *thermodynamic potentials* are functions that result from a transformation of the internal energy function $U$ from one set of variables to another.

**Terminology and Notation:**

- $P =$ pressure, $V =$ volume, $T =$ temperature($K$), $S =$ entropy
- $W =$ work done on the system, $Q =$ heat energy into the system,

**The First Law of Thermodynamics:**

$$dU = dQ + dW = TdS - PdV$$

**Internal Energy and the Thermodynamic Potentials:**

- $U =$ Internal Energy, $U = U(S,V)$, $H =$ Enthalpy, $H = H(S,P)$,
- $F =$ Helmholtz Free Energy, $F = F(T,V)$
Example 11 Demonstrate a Legendre Transformation on the First Law of Thermodynamics by deriving the form of the thermodynamic potential Helmholtz Free Energy such that $F = F(T, V)$.

Solution: Let the internal energy, $U$, is a function of entropy, $S$, and volume, $V$. Then the total differential for the internal energy can be written as

$$dU (S, V) = \left( \frac{\partial U}{\partial S} \right)_V dS + \left( \frac{\partial U}{\partial V} \right)_S dV. \quad (5.160)$$

Comparing this equation with the first low of thermodynamics

$$dU = TdS - PdV \quad (5.161)$$

we find

$$\left( \frac{\partial U}{\partial S} \right)_V = T, \quad (5.162)$$

$$\left( \frac{\partial U}{\partial V} \right)_S = -P. \quad (5.163)$$

The Helmholtz free energy which is a function of $T$ and $V$, can then be written as

$$F = F(T, V) = F \left( \left( \frac{\partial U}{\partial S} \right)_V, V \right) = F \left( \left( \frac{\partial U}{\partial S} \right)_V \right), \quad (5.164)$$

which is the Legendre Transformation of the internal energy $U$

$$F \left( \left( \frac{\partial U}{\partial S} \right)_V \right) = S \left( \frac{\partial U}{\partial S} \right)_V - U. \quad (5.165)$$

If we substitute

$$\left( \frac{\partial U}{\partial S} \right)_V = T \quad (5.166)$$

we find

$$F = F(T, V) = ST - U \quad (5.167)$$

5.8 Homework Assignment 10

1. For the function

$$z (u, v, w) = \ln \sqrt{u^2 + v^2 + w^2}$$

find the following partial derivatives:

$$\frac{\partial z}{\partial u}, \frac{\partial z}{\partial v}, \text{ and } \frac{\partial z}{\partial w}.$$
2. For the function 
\[ u(x, y) = e^x \cos(y) \]
verify that
\[ \frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x} \]
(b) 
\[ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \]

If 
\[ z(x, y) = x^2 + 2y^2, x = r \cos(\theta), y = r \sin(\theta) \]
find the following partial derivatives
3. 
\[ \left( \frac{\partial z}{\partial \theta} \right)_x \]
4. 
\[ \frac{\partial z}{\partial y \partial \theta} \]

5. Use differential to show that for large \( n \) and small \( a \)
\[ \sqrt{n + a} - \sqrt{n} \approx \frac{a}{2\sqrt{n}}. \]

5.9 Homework Assignment 11

1. The acceleration of gravity can be found from the length \( l \) and period \( T \)
of a pendulum; the formula is 
\[ g = 4\pi^2 l/T^2. \]
Find the relative error in \( g \) in the worst case if the relative error in \( l \) is 5%, and the relative error in \( T \) is 2%.

2. For an idea gas of \( N \) molecules, the number of molecules with speeds \( \leq v \)
is given by the formula
\[ n(v) = \frac{4a^3N}{\sqrt{\pi}} \int_0^v x^2 e^{-a^2 x^2} \, dx \]
where \( a \) is a constant and \( N \) is the total number of molecules. If \( N = 10^{26} \),
estimate the number of molecules with speeds between \( v = 1/a \) and \( 1.01/a \)

3. Given \( z = xe^{-y}, x = \cosh(t), y = \cos(t) \), find \( \frac{dz}{dt} \).

4. If \( xy^3 - yx^3 = 6 \) is the equation of the curve, find the slope and the equation of the tangent line at the point \((1, 2)\). Computer plot the curve and the tangent line on the same axis.

5. In problem 4 find \( d^2 y/dx^2 \) at \((1, 2)\).
Chapter 6

Introduction to Differential Calculus II

In the previous chapter we got introduced to the basics in differential calculus and its applications. Generally, we introduced to how to determine the differentiation of functions of single and multivariable variables and its application maximization or minimization in physical problems. In this chapter, we will introduce to a sort of the reverse process in differential calculus and its application in real physical problems. We will learn the different methods in differential calculus used to determine a single or multivariable variable functions satisfying a differential equation. A differential equation (DE) is an equation involving derivatives of functions with respect to one or more independent variables. In real physical problems we could find linear or none linear differential equations for scalar (e.g. an electrical potential) or vector (e.g. an electric field) functions. In real physical problems we could find linear or none linear differential equations for scalar (e.g. an electrical potential) or vector (e.g. an electric field) functions. In this chapter we will consider primarily focus only in linear differential equations.

6.1 Linear ordinary differential equations

As we just stated a differential equation is an equation involving derivatives of functions with respect to one or more independent variables. If there is only one independent variable, then the equation is said to be an ordinary differential equation (ODE). The order of the ODE or any DE is the order of the highest-order derivative in the DE. An \( n^{th} \) order ODE for the function \( y(x) \) (with \( x \) as
independent and \( y \) as dependent variable) has the generic form

\[
a_0(x)y + a_1(x)\frac{dy}{dx} + a_2(x)\frac{d^2y}{dx^2} + a_3(x)\frac{d^3y}{dx^3} + a_4(x)\frac{d^4y}{dx^4} + \ldots + a_n(x)\frac{d^n y}{dx^n} = f(x) \tag{6.1}
\]

or, more concisely,

\[
\sum_{i=0}^{n} a_i(x)\frac{d^i y}{dx^i} = f(x). \tag{6.2}
\]

Note that in ODE \( a_i(x) \) are generally function of the independent variable \( x \). When \( a_i(x) = \text{Constant} \) for all \( i \), the ODE is known as linear ordinary differential equations (LODE).

In this section we learn how can determine the solutions to LODEs. By a solution to a DE we mean a function which, when substituted back into the DE, yields an identity. The general solution to an \( n^{th} \)-order LDE contains \( n \) independent, arbitrary constants of integration. A particular solution to a DE is the general solution evaluated for particular values of its arbitrary constants. These values for the constants are determined by applying boundary conditions (BC’s) that pertain to the DE, usually obtained from an analysis of the physical system being modeled by the DE.

### 6.1.1 First-order LODE with constant coefficients

First-order LODE with constant coefficients, generally, has the form

\[
a_0y(x) + a_1\frac{dy(x)}{dx} = f(x) \tag{6.3}
\]

The solutions to first order LODE can generally be determined by a method of integration. For \( a_1 \neq 0 \), one can put Eq. \( (6.3) \) in the form

\[
\frac{a_0}{a_1}y(x) + \frac{dy(x)}{dx} = \frac{1}{a_1}f(x). \tag{6.4}
\]

Noting that

\[
e^{-\frac{a_0}{a_1}x}e^{\frac{a_0}{a_1}t} = 1 \Rightarrow e^{-\frac{a_0}{a_1}x}a_1\frac{d}{dx} e^{\frac{a_0}{a_1}t} = 1 \Rightarrow \frac{a_0}{a_1} = e^{-\frac{a_0}{a_1}x} \frac{d}{dx} e^{\frac{a_0}{a_1}t} \tag{6.5}
\]

Eq. \( (6.3) \) can be rewritten as

\[
y(x)e^{-\frac{a_0}{a_1}x}d\left[ e^{\frac{a_0}{a_1}t} \right] + \frac{dy(x)}{dx} = \frac{1}{a_1}f(x). \tag{6.6}
\]

Multiplying both sides of this equation by \( e^{\frac{a_0}{a_1}t} \), we find

\[
y(x)\frac{d}{dx} \left[ e^{\frac{a_0}{a_1}t} \right] + e^{\frac{a_0}{a_1}x}\frac{dy(x)}{dx} = \frac{1}{a_1}f(x)e^{\frac{a_0}{a_1}t} \tag{6.7}
\]
which can be put in the form
\[ \frac{d}{dx} \left[ y(x) e^{a_0 x} \right] = \frac{1}{a_1} f(x) e^{a_0 x}. \] (6.8)

Integrating Eq. (6.8), we may write,
\[ \int_{x_0}^{x} \frac{d}{dx'} \left[ y(x') e^{a_0 x'} \right] dx' = \frac{1}{a_1} \int_{x_0}^{x} f(x') e^{a_0 x'} dx'. \] (6.9)

which leads to
\[ y(x) e^{a_0 x} - y(x_0) e^{a_0 x_0} = \frac{1}{a_1} \int_{x_0}^{x} f(x) e^{a_0 x} dx \] (6.10)
\[ \Rightarrow y(x) e^{a_0 x} = y(x_0) e^{a_0 x_0} + \frac{1}{a_1} \int_{x_0}^{x} f(x') e^{a_0 x'} dx'. \]

In this equation, \( a_0 \) and \( a_1 \) are constants in the DE and \( x_0 \) and \( y(x_0) \) are also constants defined by the boundary conditions of the function \( y(x) \) that pertain to the DE. It usually obtained from an analysis of the physical system being modeled by the DE. We also note that the definite integral can be expressed as
\[ \frac{1}{a_1} \int_{x_0}^{x} f(x') e^{a_0 x'} dx' = \frac{1}{a_1} \int f(x) e^{a_0 x} dx - C(x_0), \]
where we have replaced the result of the integral when it is evaluated at \( x_0 \) by a constant \( C(x_0) \). Thus Eq. (6.10) can be expressed as
\[ y(x) e^{a_0 x} = y(x_0) e^{a_0 x_0} + \frac{1}{a_1} \int f(x) e^{a_0 x} dx. \] (6.11)

Thus by introducing another constant \( C \) for
\[ C = y(x_0) e^{a_0 x_0} - C(x_0), \] (6.12)
the general solution to the first-order LODE in Eq. (6.3) is given by
\[ y(x) = Ce^{-a_0 x} + \frac{e^{-a_0 x}}{a_1} \int f(x) e^{a_0 x} dx. \] (6.13)

In the next example we will see how we apply Eq. (6.13) to solve real physical problems described by a first-order LODE.

**Example 1** An object is dropped from rest and has a vertically downward acceleration whose magnitude is given by \( a = g \exp(-kt) \), where \( k \) is a (real) positive constant and \( g \) is the acceleration due to gravity at the earth’s surface.

(a) Find the general solution for the distance fallen as a function of time.
(b) Evaluate the general solution at the appropriate BCs in order to find the particular solution for the distance fallen by the object described above as a function of time.

(c) Comment on limiting behavior of your answer, and physically describe the system being modeled.

Solution:

(a) we recall that the acceleration, \( a \), and the velocity, \( v \), are related by the equation

\[
a = \frac{dv}{dt} \tag{6.14}
\]

and using the expression for the given acceleration, we find

\[
\frac{dv}{dt} = g \exp (-kt). \tag{6.15}
\]

In view of Eqs. (6.3) and (6.13), for a LODE of the form

\[
a_0 v(t) + a_1 \frac{dv(t)}{dt} = f(t) \tag{6.16}
\]

the solution can be expressed as

\[
v(t) = C_1 e^{-\frac{a_0}{a_1} t} + \frac{e^{-\frac{a_0}{a_1} t}}{a_1} \int_{t_0}^{t} f(t') e^{\frac{a_0}{a_1} t'} dt', \tag{6.17}
\]

From Eqs. (6.15) and (?), we note that

\[a_0 = 0, a_1 = 1, f(t) = g \exp (-kt)\]

and the solution becomes

\[
v(t) = C_1 + \int_{0}^{t} g \exp (-kt') dt', \tag{6.18}
\]

where we set the initial time \( t_0 = 0 \). Upon carrying out the integration, we find

\[
v(t) = C_1 + \frac{g}{k} (1 - \exp (-kt)) \tag{6.19}
\]

To find the position we note that the velocity is related to the displacement by

\[
v(t) = \frac{dy}{dt} \Rightarrow \frac{dy(t)}{dt} = C_1 + \frac{g}{k} (1 - \exp (-kt)) \tag{6.20}
\]

Noting that this equation can be put

\[
b_0 y(t) + b_1 \frac{dy(t)}{dt} = f(t) \tag{6.21}
\]
where
\[ b_0 = 0, b_1 = 1, f(t) = C_1 + \frac{g}{k} - \frac{g}{k} \exp(-kt) \] (6.22)

Applying Eq. (6.17) we can write the solution to Eq. (6.17) as
\[ y(t) = C_2 e^{\frac{b_0}{b_1}t} + \frac{e^{\frac{b_0}{b_1}t}}{b_1} \int_{t_0}^{t} f(t') e^{\frac{-b_0}{b_1}t'} dt', \] (6.23)

so that substituting the values in Eq. (6.22), we have
\[ y(t) = C_2 + \int_{t_0}^{t} \left[ C_1 + \frac{g}{k} - \frac{g}{k} \exp(-kt') \right] dt', \] (6.24)

and upon carrying out the integration, one finds
\[ y(t) = C_2 + \left( C_1 + \frac{g}{k} \right) t + \frac{g}{k} [\exp(-kt) - 1] \] (6.25)

(b) We are given the initial conditions (boundary conditions) \( v(t = 0) = 0 \)
and assume that the object is dropped from an initial position \( y(t = 0) = y_0 \).
Then evaluating Eqs. (6.19) and (6.25) at \( t = 0 \), we find
\[ C_1 = 0, C_2 = y_0 \]
and the distance fallen as function of time becomes
\[ d(t) = y_0 - y(t) = \frac{g}{k} t + \frac{g}{k^2} [1 - \exp(-kt)] \] (6.26)

and the speed
\[ v(t) = \frac{g}{k} (1 - \exp(-kt)) \] (6.27)

(c) For small \( t \) (using Taylor series expansion), we have
\[ \exp(-kt) \simeq 1 - kt + \frac{1}{2} kt^2 \Rightarrow 1 - \exp(-kt) \simeq kt - \frac{1}{2} kt^2 \] (6.28)
so that
\[ d(t) \simeq \frac{g}{k} t + \frac{g}{k^2} \left[ kt - \frac{1}{2} kt^2 \right] \Rightarrow d(t) \simeq \frac{1}{2} gt^2. \] (6.29)

and the speed
\[ v(t) \simeq \frac{g}{k} (kt - \frac{1}{2} kt^2) \Rightarrow v(t) \simeq gt \] (6.30)
where we have dropped the higher order in \( t \). For large time, the velocity
\[ v(t) = \frac{g}{k} (1 - \exp(-kt)) \] (6.31)
becomes
\[ v(t) \simeq \frac{g}{k} = v_0. \] (6.32)
For the distance
\[ d(t) = y_0 - y(t) = \frac{g}{k} t + \frac{g}{k^2} \left[ 1 - \exp(-kt) \right], \] (6.33)
if we let the large time be \( t = t_0 \), we find
\[ d(t) = d(t) = y_0 - y(t) = v_0 t_0 + \frac{g}{k^2}. \] (6.34)

**Example 2** A conducting bar illustrated in the figure below moves on two frictionless parallel semi-infinite rails connected by a conducting wire of length \( l \) in the presence of uniform magnetic field directed into the page. The bar has mass \( m \) and its length is also \( l \). The bar is given an initial velocity, \( v_0 \), at \( t = 0 \) directed to the right.

(a) Using the differential form of Faraday’s Law of induction find the current induced in the circuit assuming that the resistance of the circuit is \( R \).

(b) Find the magnetic force on the conducting bar.

(c) Using Newton’s second law find the speed of the bar as function of time.

(d) Show that the energy dissipated in the circuit (Joule Heat energy) is equal to the initial kinetic energy of the bar.

**Solution:**

(a) According to Faraday’s law a changing magnetic flux, \( \phi \), induces voltage, \( V_{in} \), given by
\[ V_{in} = \frac{d\phi}{dt}. \] (6.35)
Over a time interval \( t = t \) and \( t = t + dt \), the position of the conducting rod changes from \( x_0 \) to \( x_0 + dx \), as the as it slides along the positive x-direction with a speed \( v(t) \). This distance can expressed as
\[ dx = vv(t) \, dt. \] (6.36)
In the process the area bounded by part of the rails, the wire, and the rod increases from an area \( A = x_l \) at \( t = t \), to an area, \( A = x_l + ldx \). The increase in area, \( dA \), will then be
\[
dA = ldx = lvdt. \tag{6.37}
\]
The magnetic flux, \( \phi \), for a uniform magnetic field \( B \) directed perpendicular to the area, \( A \), bounded a closed conducting wire can be expressed as
\[
\phi = BA. \tag{6.38}
\]
Since the change in the magnetic flux with time is due to the change in area, we can write
\[
d\phi = BdA. \tag{6.39}
\]
so that using Eq. (6.37) one finds for the change flux over the time interval \( dt \)
\[
d\phi = Blvdt. \tag{6.40}
\]
The induced voltage due to the change in magnetic flux with time becomes
\[
V_{in} = \frac{d\phi}{dt} = Blv. \tag{6.41}
\]
Applying Ohm’s law the resulting induced current, \( I_{in} \) can be expressed as
\[
I_{in} = \frac{V_{in}}{R} = \frac{Blv}{R}. \tag{6.42}
\]
According to Lenz’s law the induced current must oppose the cause for the induction of the current. The cause is increase in magnetic flux. To oppose this increase, the induced current must flow in a counterclockwise direction.

(b) The bar experiences a magnetic force as a result of the induced current and the uniform magnetic field. This magnitude of this force is given by
\[
F_m = I_{in}lB = \frac{l^2vB^2}{R}. \tag{6.43}
\]
The direction is opposite to the direction of the velocity as one easily determine using the right-hand-rule.

c) Using Newton’s second law
\[
m \frac{dv}{dt} = -F_m = - \frac{l^2vB^2}{R} \Rightarrow \frac{l^2B^2}{mR}v + \frac{dv}{dt} = 0. \tag{6.44}
\]
that can be put in the form
\[
a_0v(t) + a_1 \frac{dv}{dt} = f(t) \tag{6.45}
\]
where

\[ a_0 = \frac{l^2 B^2}{mR}, a_1 = 1, f(t) = 0 \]

We recall the solution to Eq. (6.45) is given by

\[ v(t) = C_1 e^{\frac{a_0}{a_1} t} + \frac{e^{\frac{a_0}{a_1} t}}{a_1} \int_{t_0}^{t} f(t') e^{\frac{a_0}{a_1} t'} dt', \quad (6.46) \]

which gives

\[ v(t) = C_1 e^{-\frac{a_0}{a_1} t} = C_1 e^{-\frac{l^2 B^2 t}{mR}} \quad (6.47) \]

Noting that initial velocity, \( v(t = 0) = v_0 \), the velocity of the rod becomes

\[ v(t) = v_0 e^{-\frac{l^2 B^2 t}{mR}} ; \quad (6.48) \]

(d) The energy delivered to the resistor per unit time (power) is given by

\[ P = \frac{dW}{dt} = I_{in}^2(t) R \quad (6.49) \]

Using the results we found for the induced current in Eq. (6.42) along with the velocity in Eq. (6.48), Eq. (6.42) can be put in the form

\[ \frac{dW}{dt} = \frac{l^2 v^2(t) B^2}{R^2} R = \frac{B^2 l^2 v_0^2}{R} e^{-\frac{2l^2 B^2 t}{mR}} \quad (6.50) \]

or

\[ \frac{dW}{dt} = f(t) \quad . \quad (6.51) \]

Then the total energy delivered to the resistor is

\[ W = \frac{B^2 l^2 v_0^2}{R} \int_{0}^{\infty} e^{-\frac{2l^2 B^2 t}{mR}} dt \Rightarrow W = \frac{B^2 l^2 v_0^2}{R} \cdot \frac{mR}{2l^2 B^2} \]

\[ \Rightarrow W = \frac{1}{2} m\nu^2_0 \quad (6.52) \]

Example 3 Discuss the variation in the intensity of radiation as it passes through a stellar atmosphere assuming a plane-parallel model for the atmosphere.

Starting Point: Consider a differential slab of stellar atmosphere of thickness \( dz \) at the position \( z \) above the reference position \( z = 0 \). Let \( \rho(z) \) be the density of the atmosphere at that position, and let \( I(z) \) be the intensity of radiation incident on the slab. Let \( I_o \) be the intensity of radiation at the reference position \( z = 0 \) within the atmosphere. Find an expression for \( I(z) \) for \( z > 0 \).
6.2. The First-Order DE and Exact Differential

Solution: The change in intensity due to absorption is related to the density of the atmosphere by

\[ -dI(z) = K\rho(z)dz \]  \hspace{1cm} (6.53)

where \( K \) is a constant of proportionality.

\[
\int_{I_0}^{I(z)} dI(z) = - \int_0^z K\rho(z)dz \Rightarrow I(z) = I_0 - \int_0^z K\rho(z)dz \hspace{1cm} (6.54)
\]

6.2 The first-order DE and exact differential

In the previous section, we have seen how the solution to a first-order LODE

\[ a_0(x) y(x) + a_1(x) \frac{dy(x)}{dx} = f(x) \]  \hspace{1cm} (6.55)

can be determined when the coefficients are constant. Generally, a first-order DE could be ODE where the coefficients could be function of one variable (only \( x \)) as described by Eq. (6.55). It could also be a DE in which the coefficients are a function of two variables \( (x, y) \) that has the form

\[ Q(x,y) \frac{dy(x)}{dx} = -P(x, y). \]  \hspace{1cm} (6.56)

In this section we will develop the method to solve these two type of first-order DEs. But we first consider the later case. To this end, we note that Eq. (6.56) can be rewritten as

\[ dF(x,y) = P(x,y)dx + Q(x, y)dy = 0, \]  \hspace{1cm} (6.57)

which is the total differential of some constant function, \( F(x,y) = \text{constant} \), that depends on both \( x \) and \( y \) if

\[ P(x,y) = \frac{\partial F(x,y)}{\partial x}, Q(x,y) = \frac{\partial F(x,y)}{\partial y}. \]  \hspace{1cm} (6.58)

This indicates if one can determine this constant function from the total differential obtained from the first-order DE in Eq. (6.56), one can solve the DE by finding \( y \) from the function. This depends on certain conditions that the total deferential needs to satisfy that we will discuss. But first we will see how to determine the total differential, generally, for any function \( F(x,y) \), that is not necessarily a constant function.

Example 4 Find \( dF \) for the function

\[ F(x,y) = 3x^2 + 2x \sin(2y) \].  \hspace{1cm} (6.59)
CHAPTER 6. INTRODUCTION TO DIFFERENTIAL CALCULUS II

Solution: The total differential, \( dF \), for the function of two variables, \( F(x, y) \), is given by

\[
dF = P(x, y) \, dx + Q(x, y) \, dy.
\]  

(6.60)

For the given function in Eq. (6.59), we have

\[
P(x, y) = \frac{\partial F}{\partial x} = \frac{\partial}{\partial x} \left[ 3x^2 + 2x \sin(2y) \right] = 6x + 2 \sin(2y) \tag{6.61}
\]

and

\[
Q(x, y) = \frac{\partial F}{\partial y} = \frac{\partial}{\partial y} \left[ 3x^2 + 2x \sin(2y) \right] = 4x \cos(2y) \tag{6.62}
\]

so that total differential becomes

\[
dF = [6x + 2 \sin(2y)] \, dx + [4x \cos(2y)] \, dy \tag{6.63}
\]

In Example 4 we found the total differential

\[
dF = P(x, y) \, dx + Q(x, y) \, dy, \tag{6.64}
\]

where

\[
P(x, y) = \frac{\partial F}{\partial x} = 6x + 2 \sin(2y) \quad \text{and} \quad Q(x, y) = \frac{\partial F}{\partial y} = 4x \cos(2y). \tag{6.65}
\]

Taking the partial derivative of the function \( P(x, y) \) with respect to \( y \) and the function \( Q(x, y) \) with respect to \( x \), we find

\[
\frac{\partial P}{\partial y} = \frac{\partial F}{\partial y \partial x} = \frac{\partial}{\partial y} [6x + 2 \sin(2y)] = 4 \cos(2y) \tag{6.66}
\]

\[
\frac{\partial Q}{\partial x} = \frac{\partial F}{\partial x \partial y} = \frac{\partial}{\partial x} [4x \cos(2y)] = 4 \cos(2y). \tag{6.67}
\]

This shows that for the total differential

\[
dF = P(x, y) \, dx + Q(x, y) \, dy, \tag{6.68}
\]

we found

\[
\frac{\partial P(x, y)}{\partial y} = \frac{\partial Q(x, y)}{\partial x}. \tag{6.69}
\]

Such kind of differential is called an exact differential. If the differential is an exact differential, then there is a function \( F(x, y) \) such that

\[
P(x, y) = \frac{\partial F}{\partial x}, Q(x, y) = \frac{\partial F}{\partial y}.
\]

Example 5 For the first-order DE

\[
\left( x^3 + 2y \right) \frac{dy}{dx} = 3x (2 - xy), \tag{6.70}
\]
6.2. THE FIRST-ORDER DE AND EXACT DIFFERENTIAL

(a) Find the total differential for a function of two variables \( F(x, y) = \text{constant} \) in form

\[
dF(x, y) = P(x, y) \, dx + Q(x, y) \, dy = 0 \quad (6.71)
\]

(b) Show that the total differential is an exact differential and find the function \( F(x, y) = \text{constant} \).

(c) Using the function \( F(x, y) = \text{constant} \), find the solution to the given differential equation.

Solution:

(a) Multiplying the given first-order differential equation by \( dx \), we find

\[
(x^3 + 2y) \, dy = 3x(2 - xy) \, dx. \quad (6.72)
\]

This can be put in the form

\[
dF(x, y) = P(x, y) \, dx + Q(x, y) \, dy = 0, \quad (6.73)
\]

where

\[
P(x, y) = 3x(2 - xy), \quad Q(x, y) = -(x^3 + 2y). \quad (6.74)
\]

(b) Now taking the partial derivative of \( P \) with respect to \( y \) and \( Q \) with respect to \( x \), we find

\[
\frac{\partial P}{\partial y} = \frac{\partial}{\partial y} [3x(2 - xy)] = -3x^2 \quad (6.75)
\]

and

\[
\frac{\partial Q}{\partial x} = \frac{\partial}{\partial x} [- (x^3 + 2y)] = -3x^2. \quad (6.76)
\]

There follows that

\[
\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad (6.77)
\]

Thus, the total differential is an exact differential and there exists a function \( F(x, y) = \text{constant} \), such that

\[
P = \frac{\partial F}{\partial x} \quad \text{and} \quad Q = \frac{\partial F}{\partial y}. \quad (6.78)
\]

Now using

\[
P = \frac{\partial F}{\partial x} = 3x(2 - xy) \quad (6.79)
\]

and integrating with respect to \( x \)

\[
\int_{x_0}^{x} \frac{\partial F(x', y)}{\partial x'} \, dx' = \int_{x_0}^{x} 3x'(2 - x'y) \, dx'
\]

we find

\[
F(x, y) - F(x_0, y) = 3x^2 - x^3y - 3x_0^2 + x_0^3y. \quad (6.81)
\]
This equation can be rewritten as

\[ F(x, y) = 3x^2 - x^3y + C(y), \quad (6.82) \]

where we introduced the function

\[ C(y) = F(x_0, y) - 3x_0^2 + x_0^3y, \quad (6.83) \]

is a constant with respect to \( x \) but not with respect to \( y \). Now substituting Eq. (6.82) into

\[ Q(x, y) = \frac{\partial F}{\partial y} = -(x^3 + 2y), \quad (6.84) \]

we find

\[ \frac{\partial}{\partial y} [3x^2 - x^3y + C(y)] = -(x^3 + 2y) \]

\[ \Rightarrow -x^3 + \frac{dC(y)}{dy} = -x^3 - 2y \Rightarrow \frac{dC(y)}{dy} = -2y \quad (6.85) \]

Now integrating with respect to \( y \),

\[ \int_{y_0}^{y} \frac{dC(y')}{dy'} dy' = \int_{y_0}^{y} y' dy' = \int_{y_0}^{y} y dy \quad (6.86) \]

we find

\[ C(y) = -y^2 + C_1, \quad (6.87) \]

where

\[ C_1 = y_0^2 + C(y_0), \]

is constant independent of \( x \) and \( y \). Therefore, the solution to the DE is found to be

\[ F(x, y) = 3x^2 - x^3y + C(y) = 3x^2 - x^3y - y^2 = C. \quad (6.88) \]

(c) The equation

\[ 3x^2 - x^3y - y^2 = C \quad (6.89) \]

can be rewritten as

\[ y^2 + x^3y - 3x^2 + C = 0, \quad (6.90) \]

or

\[ ay^2 + by + c = 0, \quad (6.91) \]

which is a quadratic equation with

\[ a = 1, b = x^3, c = -3x^2 + C. \quad (6.92) \]

Using the well known solutions for a quadratic equation,

\[ y = \frac{b \pm \sqrt{b^2 - 4ac}}{2a} \quad (6.93) \]

we find

\[ y(x) = \frac{1}{2} \left[ x^3 \pm \sqrt{x^6 + 4(3x^2 - C)} \right] \quad (6.94) \]
Example 6 Solve the DE
\[(y^2 - 3x^2 e^{3y}) \frac{dy}{dx} = 2xe^{3y} + e^x. \quad (6.95)\]
in the same ways as Example 6.

Solution:
(a) Multiplying the given first-order differential equation by \(dx\), we find
\[(y^2 - 3x^2 e^{3y}) dy = (2xe^{3y} + e^x) dx. \quad (6.96)\]
or
\[dF(x, y) = P(x, y) dx + Q(x, y) dy = 0, \quad (6.97)\]
where
\[P(x, y) = 2xe^{3y} + e^x, \quad Q(x, y) = -(y^2 - 3x^2 e^{3y}) \quad (6.98)\]

(b) For an exact differential, we must be able to show that
\[\frac{\partial P(x, y)}{\partial y} = \frac{\partial Q(x, y)}{\partial x}. \quad (6.99)\]
Using Eq. (6.98), we find
\[\frac{\partial P(x, y)}{\partial y} = \frac{\partial}{\partial y} (2xe^{3y} + e^x) = 6xe^{3y},\]
\[\frac{\partial Q(x, y)}{\partial x} = -\frac{\partial}{\partial x} (y^2 - 3x^2 e^{3y}) = 6xe^{3y}, \quad (6.100)\]
that shows
\[\frac{\partial P(x, y)}{\partial y} = \frac{\partial Q(x, y)}{\partial x}. \quad (6.101)\]
Therefore, the differential is an exact differential and there exists a function \(F(x, y) (=\text{Constant}, \text{in this case})\) such that
\[P(x, y) = \frac{\partial F}{\partial x}, \quad Q(x, y) = \frac{\partial F}{\partial y}. \quad (6.102)\]
This leads to
\[\int_{x_0}^{x} \frac{\partial F}{\partial x'} dx' = \int_{x_0}^{x} \left(2x'e^{3y} + x'e^x\right) dx',\]
\[\Rightarrow F(x, y) - F(x_0, y) = x^2 e^{3y} + xe^x - \left(x_0^2 e^{3y} + x_0 e^{x_0}\right) + C_1(y) \quad (6.103)\]
or
\[F(x, y) = x^2 e^{3y} + xe^x + C_1(y). \quad (6.104)\]
where we introduced the function defined by
\[F(x, y) = x^2 e^{3y} + xe^x + C_1(y). \quad (6.104)\]
Substituting this result into
\[ \frac{\partial F}{\partial y} = Q(x, y) = -(y^2 - 3x^2 e^{3y}) = 3x^2 e^{3y} - y^2, \]  
we find
\[ \frac{\partial}{\partial y} \left[ x^2 e^{3y} + xe^x + C_1(y) \right] = 3x^2 e^{3y} - y^2 \Rightarrow \frac{dC_1(y)}{dy} = -y^2. \]  
Integrating this equation with respect to \( y \) yields
\[ C_1(y) = -\int y^2 dy + K \Rightarrow C_1(y) = -\frac{y^3}{3} + C_2, \]
where \( C_2 \) is a constant of integration with respect to \( y \) and it is independent of both \( x \) and \( y \). Now substituting this result into Eq. (6.104), the solution to the exact differential in Eq. (6.104) is found to be
\[ F(x, y) = x^2 e^{3y} + xe^x - \frac{1}{3} y^3 = C. \]  
Note that we have defined \( C \) representing all the constants independent of the both variables \( x \) and \( y \).

### 6.3 First-order DE and none exact total differential

We recall that the differential equation
\[ dF = P(x, y) \, dx + Q(x, y) \, dy \]  
is said to be an exact differential equation, when
\[ \frac{\partial P(x, y)}{\partial y} = \frac{\partial Q(x, y)}{\partial x}. \]  
For an exact differential there exists a function, \( F(x, y) \), such that
\[ P(x, y) = \frac{\partial F(x, y)}{\partial x}, Q(x, y) = \frac{\partial F(x, y)}{\partial y} \Rightarrow \frac{\partial^2 F(x, y)}{\partial y \partial x} = \frac{\partial^2 F(x, y)}{\partial x \partial y}. \]  
Sometimes, we may find a first order LDE with none constant coefficients that may lead to a total differential that is not exact. Let’s consider the first-order ODE with none constant function
\[ a_0(x) y(x) + a_1(x) \frac{dy(x)}{dx} = f(x). \]
6.3. **FIRST-ORDER DE AND NONE EXACT TOTAL DIFFERENTIAL**

This ODE can be rewritten as

\[
\frac{dy(x)}{dx} = - \left( \frac{a_0(x)}{a_1(x)} y(x) - \frac{f(x)}{a_1(x)} \right)
\]  

(6.112)

that can be put in the form

\[
Q(x,y) \frac{dy(x)}{dx} = -P(x,y).
\]  

(6.113)

There follows the total deferential

\[
dF(x,y) = P(x,y) \, dx + Q(x,y) \, dy = 0,
\]  

(6.114)

where

\[
P(x,y) = \frac{a_0(x)}{a_1(x)} y(x) - \frac{f(x)}{a_1(x)}, \quad Q(x,y) = 1.
\]  

(6.115)

Here we note that

\[
\frac{\partial P(x,y)}{\partial y} = \frac{\partial}{\partial y} \left[ \frac{a_0(x)}{a_1(x)} y(x) - \frac{f(x)}{a_1(x)} \right] = y, \quad \frac{\partial Q(x,y)}{\partial x} = 0,
\]  

(6.116)

and

\[
\frac{\partial P(x,y)}{\partial y} \neq \frac{\partial Q(x,y)}{\partial x}
\]  

(6.117)

that mean the resulting DE from the first-order ODE is not an exact differential and the method we developed in the previous section does not work. However, this method can be used if we can find a function that can make the ODE in Eq. (6.111) to lead to an exact differential. Next we will determine this function which is referred as "The integration factor".

**The Integration Factor:** The first-order ODE with none constant coefficient in Eq. (6.111) can put in the form

\[
\frac{dy}{dx} + P(x)y = Q(x),
\]  

(6.118)

where we introduced the functions defined by

\[
P(x) = \frac{a_0(x)}{a_1(x)}, \quad Q(x) = \frac{f(x)}{a_1(x)}.
\]  

(6.119)

This is the generic first-order ODE. Let's say this ODE does not lead to an exact differential. Let the function \( V(x) \) be the function that when it multiplied to the ODE in Eq. (6.118) leads to a total differential that is an exact differential. Thus multiplying Eq. (6.118) by \( V(x) \), we have

\[
V(x) \frac{dy}{dx} + V(x) P(x)y = V(x) Q(x)
\]

\[
\Rightarrow V(x) \frac{dy}{dx} + V(x) P(x)y - V(x) Q(x) = 0
\]  

(6.120)
which can put in the form

\[
dF(x,y) = [V(x)P(x)y - V(x)Q(x)] \, dx + V(x) \, dy = 0.
\] (6.121)

We recall that for differential to be an exact differential the partial derivative of the coefficient of \(dx\) with respect to \(y\) and the partial derivative of the coefficient of \(dy\) with respect to \(x\) must be equal. This means for the differential in Eq. (6.121), we must have

\[
\frac{\partial}{\partial y} [V(x)P(x)y - V(x)Q(x)] = \frac{\partial}{\partial x} V(x)
\] (6.122)

that leads

\[
V(x)P(x) = \frac{dV(x)}{dx} \Rightarrow P(x)dx = \frac{dV(x)}{V(x)}.
\] (6.123)

Upon integrating this equation one finds

\[
V(x) = e^{\int P(x)dx}
\] (6.124)

The function \(V(x)\) given by the integral expression Eq. (6.124) is the multiplying factor to the ODE in Eq. (6.118) that leads to an exact differential. It is referred as the integration factor. Thus multiplying Eq. (6.118) by the integration factor, we find

\[
e^{\int P(x)dx} \frac{dy}{dx} + e^{\int P(x)dx} P(x)y = e^{\int P(x)dx} Q(x).
\] (6.125)

Applying the product rule for differentiation, we put this equation in the form

\[
\frac{d}{dx} \left[ ye^{\int P(x)dx} \right] = e^{\int P(x)dx} Q(x),
\] (6.126)

so that integrating it with respect to \(x\), one can write

\[
ye^{\int P(x)dx} = \int \left[ e^{\int P(x)dx} Q(x) \right] \, dx + C,
\] (6.127)

where \(C\) is the constant of integration. Dividing this equation by the integration factor, we find

\[
y(x) = e^{-\int P(x)dx} \int \left[ e^{\int P(x)dx} Q(x) \right] \, dx + Ce^{-\int P(x)dx}.
\] (6.128)

This is the general solution to the first-order ODE in Eq. (6.118) and the constant \(C\) is determined from the given boundary condition set to the specific problem that the DE is representing. In the next examples, we will see its application. But let’s examine Eq. (6.128) if it gives the general solution we determined for a first-order ODE with constant coefficients. We recall, for a ODE with constant coefficient

\[
a_0y(x) + a_1 \frac{dy(x)}{dt} = f(x)
\] (6.129)
6.3. FIRST-ORDER DE AND NONE EXACT TOTAL DIFFERENTIAL

the solution were shown to be

\[ y(x) = Ce^{-\frac{a_0}{a_1} x} + e^{-\frac{a_0}{a_1} x} \int \frac{f(x)}{a_1} e^{\frac{a_0}{a_1} x} \, dx. \]  

(6.130)

Noting that Eq. (6.129) can rewritten as Eq. (6.118)

\[ \frac{dy}{dx} + P(x)y = Q(x), \]  

(6.131)

where

\[ P(x) = \frac{a_0}{a_1}, \quad Q(x) = \frac{f(x)}{a_1}. \]  

(6.132)

Substituting this functions into Eq. (6.128), we find

\[ y(x) = e^{-\int \frac{a_0}{a_1} \, dx} \int \left[ e^{\int \frac{a_0}{a_1} \, dx} \frac{f(x)}{a_1} \right] \, dx + Ce^{-\int \frac{a_0}{a_1} \, dx}. \]  

(6.133)

that leads to

\[ y(x) = Ce^{-\frac{a_0}{a_1} x} + \frac{e^{-\frac{a_0}{a_1} x}}{a_1} \int e^{\frac{a_0}{a_1} x} f(x) \, dx. \]  

(6.134)

which the same as Eq. (6.130).

**Example 7** For the first-order DE

\[ 2x \frac{dy}{dx} + 4x^2 = y \]  

(6.135)

(a) Express the DE in the form

\[ \frac{dy}{dx} + P(x)y = Q(x), \]  

(6.136)

(b) Using the general solution for such kind of first-order DE (Eq. (6.118)), find the solution for \( y(x) \) subject to the boundary condition,

\[ y(x_0 = 1) = \frac{26}{3}. \]  

(6.137)

**Solution:**

(a) Noting that the DE can be rewritten as

\[ \frac{dy}{dx} + 2x = \frac{1}{2x} y \Rightarrow \frac{dy}{dx} + P(x) y = Q(x), \]  

(6.138)

we have

\[ P(x) = -\frac{1}{2x}, \quad Q(x) = -2x. \]  

(6.139)
(b) Using the function $P(x)$, we find
\[
\int P(x)dx = -\int \frac{1}{2x} dx = -\frac{1}{2} \ln x.
\]
Substituting this result and $Q(x) = -2x$ into Eq. (6.118), we have
\[
y(x) = e^{\frac{1}{2} \ln x} \int \left[ e^{-\frac{1}{2} \ln x} (-2x) \right] dx + Ce^{\frac{1}{2} \ln x} \\
= -2e^{\ln \sqrt{x}} \int e^{\ln(1/\sqrt{x})} x dx + Ce^{\ln \sqrt{x}}
\]
so that
\[
y(x) = -2\sqrt{x} \int \frac{1}{\sqrt{x}} x dx + C\sqrt{x} \Rightarrow y(x) = -\frac{4}{3} x^2 + C\sqrt{x},
\]
where we used the relations
\[
e^{\ln \sqrt{x}} = \sqrt{x} \text{ and } e^{\ln(1/\sqrt{x})} = \frac{1}{\sqrt{x}}.
\]
Using the given boundary condition
\[
y(x = 1) = \frac{26}{3}
\]
the constant of integration, $C$, is found to be
\[
-\frac{4}{3} + C = \frac{26}{3} \Rightarrow C = 10.
\]
Therefore, the general solution in Eq. (6.142) becomes
\[
y(x) = \left(10 - \frac{4}{3} e^{3/2}\right) \sqrt{x}.
\]
This is the solution to the ODE in Eq. (6.135) under the given boundary condition.

6.4 Higher-Order ODE’s

In the previous section we have introduced to various methods for solving a first-order ODE. Next we will get introduced to how to solve higher-order $(n \geq 2)$ ODEs. We recall that an $n^{th}$ order ODE is given by
\[
a_0 (x) y + a_1 (x) \frac{dy}{dx} + a_2 (x) \frac{d^2y}{dx^2} + a_3 (x) \frac{d^3y}{dx^3} \ldots a_{n-1} (x) \frac{d^{n-1}y}{dx^{n-1}} + a_n (x) \frac{d^ny}{dx^n} = f(x).
\]
We will develop the methods for solving this equation in two parts. In the first part we will consider various techniques for Homogenous ODE. Generally, a Homogeneous DE is a DE in which the function, $f(x) = 0$. In the second part, we will develop the techniques for none-Homogeneous ordinary differential equations (HODE) that will be built upon the techniques for HODE.
6.4. HIGHER-ORDER ODE’S

6.4.1 Homogeneous ODE’s with constant coefficient

From Eq. (6.146) an $n^{th}$ order HODE when all the coefficients, $a_0(x) = a_0, a_1(x) = a_1, a_2(x) = a_2, \ldots, a_n(x) = a_n$, can be written as

$$a_0y + a_1 \frac{dy}{dx} + a_2 \frac{d^2y}{dx^2} + a_3 \frac{d^3y}{dx^3} + \ldots + a_{n-1} \frac{d^{n-1}y}{dx^{n-1}} + a_n \frac{d^n y}{dx^n} = 0.$$  (6.147)

We will introduce how to determine the general solution to such kind of HODE by solving what is known as the *indicial equation* that can be derived from the HODE.

**Indicial Equation:** The indicial equation is founded on an educated guess. We guess the solution to Eq. (6.147) be

$$y(x) = Ae^{kx}. \quad (6.148)$$

Substituting this into Eq. (6.147), one can easily find

$$(a_0 + a_1 k + a_2 k^2 + a_3 k^3 + \ldots + a_{n-1} k^{n-1} + a_n k^n) Ae^{kx} = 0. \quad (6.149)$$

There follows that

$$(a_0 + a_1 k + a_2 k^2 + a_3 k^3 + \ldots + a_{n-1} k^{n-1} + a_n k^n) = 0. \quad (6.150)$$

This equation is called the *indicial equation*. The solution to the indicial equation in Eq. (6.150) gives $n$ roots that could be real or complex. If these roots are *none degenerate* (i.e. no repeated identical roots) and are given by

$$k = k_1, k_2, \ldots, k_n, \quad (6.151)$$

the general solution to the HODE with constant coefficients in Eq. (6.147) is given by

$$y(x) = C_1 e^{k_1 x} + C_2 e^{k_2 x} + \ldots + C_n e^{k_n x}. \quad (6.152)$$

As we have seen in the in the solutions to first-order ODE, the constants $C_1, C_2, \ldots, C_n$, are determined from the boundary conditions pertinent to the particular problem.

**Example 8** An inductor and capacitor are connected in a series circuit, as shown. At $t = 0$, the charge on the capacitor plates is $Q_o$ and the current flowing in the circuit is $I_o$. Find an expression for the charge on the capacitor, $Q(t)$.

**Solution:** Using Kirchoff’s rule we may write

$$\frac{Q(t)}{C} + L \frac{dI(t)}{dt} = 0. \quad (6.153)$$
and recalling that \( I(t) = \frac{dQ}{dt} \), we may write

\[
\frac{d^2Q}{dt^2} + \frac{1}{LC}Q = 0. \tag{6.154}
\]

Eq. (6.154) is a HODE with constant coefficients. Thus, we can assume a solution of the form

\[
Q(t) = A e^{kt}. \tag{6.155}
\]

Substituting this into Eq. (6.154), we find

\[
\frac{d^2}{dt^2} [A e^{kt}] + \omega^2 A e^{kt} = 0 \Rightarrow k^2 + \omega^2 = 0 \tag{6.156}
\]

the solution of which is given by

\[
k_1 = i\omega, k_2 = -i\omega \tag{6.157}
\]

where we introduced the constant defined by

\[
\omega = \frac{1}{\sqrt{LC}}. \tag{6.158}
\]

The general solution can then be expressed as

\[
Q(t) = A \exp(i\omega t) + B \exp(-i\omega t). \tag{6.159}
\]

At \( t = 0 \) the capacitor is fully charged and the charge is \( Q_0 \), this means

\[
Q(0) = A \exp(i\omega t) + B \exp(-i\omega t)|_{t=0} = Q_0 \Rightarrow A + B = Q_0. \tag{6.160}
\]

We are also given that the current at \( t = 0 \) is \( I_0 \). Thus using the definition of an electrical current, we may write

\[
I(t = 0) = \frac{dQ}{dt} \bigg|_{t=0} = I_0
\]

\[
\Rightarrow \frac{d}{dt} [A \exp(i\omega t) + B \exp(-i\omega t)] \bigg|_{t=0} = I_0 \Rightarrow A - B = \frac{I_0}{i\omega} \tag{6.161}
\]
Combining Eqs. (6.160) and (6.161), one can easily show that

\[ A = \frac{1}{2} \left[ Q_0 + \frac{I_0}{i\omega} \right], \quad B = \frac{1}{2} \left[ Q_0 - \frac{I_0}{i\omega} \right]. \tag{6.162} \]

Therefore, the charge at a given time \( t \) can be expressed as

\[ Q(t) = \frac{1}{2} \left[ Q_0 + \frac{I_0}{i\omega} \right] \exp(i\omega t) + \frac{1}{2} \left[ Q_0 - \frac{I_0}{i\omega} \right] \exp(-i\omega t), \tag{6.163} \]

or

\[ Q(t) = Q_0 \left[ \exp(i\omega t) + \exp(-i\omega t) \right] + \frac{I_0}{\omega} \left[ \exp(i\omega t) - \exp(-i\omega t) \right]. \tag{6.164} \]

Applying Euler’s formula Eq. (6.164) can be put in the form

\[ Q(t) = A \cos(\omega t) + B \sin(\omega t), \tag{6.165} \]

where

\[ C = Q_0, \quad D = \frac{I_0}{\omega}. \tag{6.166} \]

Introducing the constants defined by

\[ Q_0 = E \sin \varphi, \quad \frac{I_0}{\omega} = E \cos \varphi, \tag{6.167} \]

Eq. (6.165) can also be expressed as

\[ Q(t) = E \left[ \sin \varphi \cos(\omega t) + \cos \varphi \sin(\omega t) \right] = E \sin(\omega t + \varphi), \tag{6.168} \]

with

\[ \tan(\varphi) = \frac{Q_0}{I_0/\omega}, \quad E = \sqrt{Q_0^2 + \left(\frac{I_0}{\omega}\right)^2}. \tag{6.169} \]

If we define another constant defined in a slightly different way

\[ Q_0 = G \cos \varphi, \quad \frac{I_0}{\omega} = G \sin \varphi, \tag{6.170} \]

one can rewrite Eq. (6.165) as

\[ Q(t) = G \cos \varphi \cos(\omega t) + G \sin \varphi \sin(\omega t) = G \cos(\omega t - \varphi), \tag{6.171} \]

where

\[ \tan(\varphi) = \frac{I_0/\omega}{Q_0}, \quad G = \sqrt{Q_0^2 + \left(\frac{I_0}{\omega}\right)^2}. \tag{6.172} \]

The current

\[ I(t) = \frac{dQ}{dt} \tag{6.173} \]

can be determined using any one of the equations for the charge in Eqs. (6.164), (6.165), (6.168), or (6.171). If we chose Eq. (6.168), we find

\[ I(t) = I_0 \cos(\omega t) - Q_0 \omega \sin(\omega t) \tag{6.174} \]

for the current in the circuit.
In Example 8, we saw that the solution to the second-order DE in Eq. (6.154) can be expressed in four different forms given by Eqs. (6.164), (6.165), (6.168), and (6.171). In view of this results, we can make two generalizations to the solution of a second-order HODE that can be put in form

\[ \frac{d^2y}{dx^2} \pm \omega^2 y = 0, \]  

(6.175)

when \( \omega \) a real constant. For the plus case,

\[ \frac{d^2y}{dx^2} + \omega^2 y = 0, \]  

(6.176)

as we saw in the example above, the roots to the indicial equation

\[ k^2 + \omega^2 = 0, \]  

(6.177)

are complex, \( k_1 = i\omega \) and \( k_2 = -i\omega \) and the general solution can be expressed in four different forms given by:

\[ y(x) = \begin{cases} 
A \exp(i\omega x) + B \exp(-i\omega x), \\
C \cos(\omega x) + D \sin(\omega x), \\
E \sin(\omega x + \varphi), \\
G \cos(\omega x + \varphi). 
\end{cases} \]  

(6.178)

For the minus case

\[ \frac{d^2y}{dx^2} - \omega^2 y = 0, \]  

(6.179)

the roots to the corresponding indicial equation

\[ k^2 - \omega^2 = 0, \]  

(6.180)

are real, \( k_1 = \omega \) and \( k_2 = -\omega \). It can easily be shown that the solution can be expressed as

\[ y(x) = \begin{cases} 
A \exp(\omega x) + B \exp(-\omega x), \\
C \cosh(\omega x) + D \sinh(\omega x), \\
E \sinh(\omega x + \varphi), \\
G \cosh(\omega x + \varphi). 
\end{cases} \]  

(6.181)

Problem 1: show the solutions in Eq. (6.181)

Problem 2: Find the current and the charge in the previous example if a resistor \( R \) is included in series with \( C \) and \( L \).

Example 10 The Schröedinger Equation: The one-dimensional time-independent Schröedinger equation for a particle traveling along the x-direction with a potential energy function \( U(x) \) is given by

\[ -\frac{\hbar^2}{2m} \frac{d^2\Psi(x)}{dx^2} + U(x)\Psi(x) = E\Psi(x), \]  

(6.182)
where $h = h/2\pi$ ($h$ is a constant called Planck’s constant), $m$ is the mass of the particle, and $E$ is the total energy of the particle. Consider a particle of mass $m$ and energy $E$ traveling in the potential shown in the figure. The x-axis is divided into two regions by the potential energy function. Region I has $U(x) = 0$, which means that there are no forces acting on the particle. In this region the particle is a free particle. Region II has $U(x) = U_o$, where $U_o$ is a constant such that $U_o > E$. This is a forbidden region for the particle according to classical physics. The significance of the wave function $\Psi(x)$ that is the solution to the Schröedinger equation above for a given potential energy function is that $|\Psi(x)|^2 \, dx$ gives the probability of finding the particle within an interval $dx$ at each value of $x$. Solve the Schröedinger equation for the wave function of the particle in each of the two regions I and II shown in the figure. Comment on your answers.

Solution: In region I, the potential energy is zero, $U(x) = 0$, and the Schröedinger Equation becomes

$$\frac{-\hbar^2}{2m} \frac{d^2\Psi(x)}{dx^2} = E\Psi(x) \Rightarrow \frac{d^2\Psi(x)}{dx^2} = -\frac{2mE}{\hbar^2} \Psi(x) \quad (6.183)$$

which we may write as

$$\frac{d^2\Psi(x)}{dx^2} = -k^2\Psi(x) = \frac{d^2\Psi(x)}{dx^2} + k^2\Psi(x) = 0 \quad (6.184)$$

where

$$k^2 = \frac{2mE}{\hbar^2}. \quad (6.185)$$

Substituting

$$\Psi(x) = Ae^{\lambda x} \quad (6.186)$$
we find the indicial equation
\[ \lambda^2 + k^2 = 0 \implies \lambda = \pm ik \] (6.187)
so that the solution can be written as
\[ \Psi(x) = A_1 e^{ikx} + A_2 e^{-ikx}. \] (6.188)

In region II, the potential energy is zero, \( U(x) = U_0 \), and the Schröedinger Equation becomes
\[
\frac{-\hbar^2}{2m} \frac{d^2 \Psi(x)}{dx^2} = -(U_0 - E) \Psi(x)
\]
\[ \Rightarrow \frac{d^2 \Psi(x)}{dx^2} = \frac{2m (U_0 - E)}{\hbar^2} \Psi(x) \] (6.189)
which we may write as
\[
\frac{d^2 \Psi(x)}{dx^2} = q^2 \Psi(x) = \frac{d^2 \Psi(x)}{dx^2} - q^2 \Psi(x) = 0
\] (6.190)
where
\[ q^2 = \frac{2m (U_0 - E)}{\hbar^2}. \] (6.191)

Substituting
\[ \Psi(x) = Ae^{\lambda x} \] (6.192)
we find the indicial equation
\[ \lambda^2 - q^2 = 0 \implies \lambda = \pm q \] (6.193)
so that the solution can be written as
\[ \Psi(x) = A_1 e^{qx} + A_2 e^{-qx}. \] (6.194)

But physical condition requires \( A_1 = 0 \), which means
\[ \Psi(x) = A_2 e^{-qx}. \] (6.195)

**Indicial equation with degenerate roots:** As we stated earlier the general solution in Eq. (??) is valid when there are no repeated roots (nondegenerate roots). However, it is possible that a HODE with constant coefficients could lead to an indicial equation with degenerate roots. For example, for the second-order HODE
\[
\frac{d^2 y}{dx^2} - 2a \frac{dy}{dx} + a^2 y = 0,
\] (6.196)
the roots of the quadratic indicial equation

\[ k^2 - 2ak + a^2 = 0, \]  

are degenerate

\[ k_1 = k_2 = a. \]  

When the roots are degenerate the general solution in Eq. (??) can be used with some changes that depend on the degeneracy in the roots of the indicial equation. For example, for the HODE in Eq. (6.196), the values for the roots are identical \((k_1 = k_2 = a)\), where the degeneracy, \(m = 2\), the general solution should be written as

\[ y(x) = (A + Bx) e^{ax}. \]  

Note that for \(m = 2\), the coefficient to \(e^{ax}\) is a polynomial function, \(P_1(x) = C_0 + C_1 x + C_2 x^2 + \ldots + C_{m-1} x^{m-1}\). Similarly, for a HODE that leads to an indicial equation, with \(m = 3\) degenerate roots, \((k_1 = k_2 = k_3 = a)\), the general solution is given by

\[ y(x) = (A + Bx + C x^2) e^{ax} = P_2(x) e^{ax}. \]  

In this case the coefficient polynomial function, \(P_2(x)\), has a degree, \(l = m - 1 = 2\).

Thus we can make the following changes to the general solution in Eq. (??) in relation to the degenerate roots to the indicial equation:

(a) \(n\)-th order differential equation leads to an indicial equation with degenerate roots, \(m = n\), the general solution is given by

\[ y(x) = (A + Bx + C x^2 + \ldots + D x^{n-1}) e^{ax} \]  

(b) \(n\)-th order differential equation that leads to an indicial equation with three different set of roots. Suppose the first set of roots has \(m = m_1\) degenerate roots (i.e. \(k_1 = k_2 = k_3 = \ldots = k_{m_1} = a\)), the second set has \(m = m_2\) degenerate roots (i.e. \(k'_1 = k'_2 = k'_3 = \ldots = k'_{m_2} = a'\)), and the third set consist of \(m_3\) none-degenerate roots \((k''_1 = a''_1, k''_2 = a''_2, k''_3 = a''_3, \ldots = k''_{m_3} = a''_{m_3})\) (Note that \(m_1 + m_2 + m_3 = n\)). For this HODE equation, the general solution is given by

\[ y(x) = P_{m_1-1}(x) e^{a'x} + P_{m_2-1}(x) e^{a''x} + \ldots + C_{m_3-1} e^{a''_{m_3}x} \]  

where

\[ P_{m_1-1}(x) = C_0 + C_1 x + C_2 x^2 + \ldots + C_{m_1-1} x^{m_1-1}, \]  

\[ P_{m_2-1}(x) = C'_0 + C'_1 x + C'_2 x^2 + \ldots + C'_{m_2-1} x^{m_2-1}, \]  

are polynomials of degree \(m_1 - 1\) and \(m_2 - 1\), respectively.
Example 11 In each of the following parts, find the indicial equation, list the roots of the indicial equation \( \{k_i\} \), and write out the general solution to the DEs:

(a) \[
\frac{d^3y}{dx^3} + \frac{dy}{dx} = 0.
\] (6.204)

(b) \[
2\frac{d^4s}{dt^4} - 3\frac{d^3s}{dt^3} = 0.
\] (6.205)

Solution:

(a) Upon substituting \( y(x) = Ae^{kx} \Rightarrow \frac{dy}{dx} = k y(x) \Rightarrow \frac{d^3y}{dx^3} = k^3 y(x) \) into the given DE, the indicial equation becomes

\[
k^3 + k = 0.
\] (6.207)

The is a cubic equation with three distinct roots

\[
k^3 + k = 0 \Rightarrow k (k^2 + 1) = 0 \Rightarrow k_1 = 0, k_2 = i, k_3 = -i.
\] (6.208)

The roots do not show any degeneracy in this case and the solution to the DE can be written as

\[
y(x) = A + Be^{ix} + Ce^{-ix}.
\] (6.209)

or

\[
y(x) = A + D \cos(x) + E \sin(x)
\] (6.210)

(b) Substituting \( s(t) = Ae^{kt} \Rightarrow \frac{d^3s}{dt^3} = k^3 s(t) \Rightarrow \frac{d^4s}{dt^4} = k^4 s(t) \) in the given DE, the roots to the indicial equation are found to be

\[
2k^4 - 3k^3 = k^3 (2k - 3) = 0 \Rightarrow k_1 = \frac{3}{2}, k_2 = k_3 = k_4 = 0.
\] (6.212)

Since the the indicial equation has three identical roots, the solution can be written as

\[
2k^4 - 3k^3 = 0 \Rightarrow k^3 (2k - 3) = 0 \Rightarrow k_1 = \frac{3}{2}, k_2 = k_3 = k_4 = 0
\] (6.213)
This shows that the three roots \((k_2, k_3, k_4)\) are degenerate with degeneracy, \(m = 3\), and one non-degenerate root \((k_1)\). Thus applying the relation in Eq. (6.202), the general solution can be written as

\[
y(x) = (A + Bx + Cx^2)e^{ax} + Ae^{k_1x},
\]

(6.214)

where \(a\) is the value for the degenerate roots and \(k_1\) is the value for the non-degenerate root. Using these values the general solution to the DE becomes

\[
y(x) = A + Bx + Cx^2 + Ae^{\frac{3}{2}x}.
\]

(6.215)

**N.B.** You may use Mathematica to check your results. Here is the result for the previous two examples.

\[
\text{In}[2] = \text{DSolve}[y''''[x] + y'[x] == 0, y, x]
\]

\[
\]

\[
\text{In}[3] = \text{DSolve}[2y''''[x] - 3y'''[x] == 0, y, x]
\]

\[
\text{Out}[3] = \{(y -> \text{Function}[[x], \frac{8}{27} e^{\frac{3}{2}x} t C[1] + C[2] + C[3] + x C[4]]\})\}
\]

Here \(C[1], C[2], C[3]\), and \(C[4]\) are constants of integration. Note that \(-1\) in the first solution can be absorbed in \(C[2]\) and \(\frac{8}{27}\) in the second solution can also be absorbed in \(C[1]\).

**Example 12** Damped Simple Harmonic Motion: An ideal spring of spring constant \(k\) hangs vertically from a support rod and has a mass \(m\) attached to its lower end. The mass hangs suspended in a viscous fluid. The mass is pulled down some distance from its equilibrium position at \(y = 0\) and released. Discuss the resulting motion of the mass if the viscous fluid exerts a damping force on it given by \(f_y = -bv_y\), where \(b\) is a damping constant.
Solution: Since the mass initially was at equilibrium and also consider the displacement from this equilibrium position, we do not consider the buoyant and gravitational force on the mass. Therefore, the equation of motion for the mass is given by Newton’s second law

\[ F_{\text{net}} = m a_y = m \frac{d^2 y}{dt^2} \quad (6.216) \]

The net force acting on the mass is given by

\[ F_{\text{net}} = -k y - b \frac{dy}{dt} = -k y - b \frac{d^2 y}{dt^2} \quad (6.217) \]

which leads to

\[ m \frac{d^2 y}{dt^2} = -k y - b \frac{dy}{dt} \Rightarrow \frac{d^2 y}{dt^2} + 2 \gamma \frac{dy}{dt} + \omega^2 y = 0 \quad (6.218) \]

where

\[ \omega^2 = \frac{k}{m}, \gamma = \frac{b}{2m}. \quad (6.219) \]

The indicial equation would then be

\[ \lambda^2 + 2 \gamma \lambda + \omega^2 = 0, \quad (6.220) \]

and the solutions

\[ \lambda_1 = -\gamma + D, \lambda_2 = -\gamma - D, \quad (6.221) \]

where

\[ D = \sqrt{\gamma^2 - \omega^2} \quad (6.222) \]

Case 1 (\( \gamma = \omega \)) we have

\[ D = \sqrt{\gamma^2 - \omega^2} = 0 \quad (6.223) \]

and the roots to the indicial equation becomes

\[ \lambda_1 = \lambda_2 = -\gamma \quad (6.224) \]

and the general solution is given by

\[ y(t) = (At + B) e^{-\gamma t} \quad (6.225) \]

This describes a **Critically damped motion**

Case 2 (\( \gamma > \omega \)) we note that \( D \) is real

\[ D = \sqrt{\gamma^2 - \omega^2} < \gamma \quad (6.226) \]

which means the solution is given by

\[ y(t) = A e^{-(\gamma - D)t} + B e^{(\gamma + D)t}. \quad (6.227) \]

This shows **over-damped motion**
6.4. HIGHER-ORDER ODE’S

Case 3 ($\gamma < \omega$) we may write
\[ D = \sqrt{\gamma^2 - \omega^2} = \sqrt{(\omega^2 - \gamma^2)} = i\Omega \] (6.228)

where
\[ \Omega = \sqrt{\omega^2 - \gamma^2} \] (6.229)

which means the roots for the indicial equation becomes
\[ \lambda_1 = -\gamma + i\Omega, \lambda_2 = -\gamma - i\Omega \] (6.230)

and the solution is given by
\[ y(t) = Ae^{-(\gamma - i\Omega)t} + Be^{-(\gamma + i\Omega)t} \Rightarrow y(t) = e^{-\gamma t} \left( Ae^{i\Omega t} + Be^{-i\Omega t} \right) \] (6.231)

or
\[ y(t) = e^{-\gamma t} [A' \cos (\Omega t) + B' \sin (\Omega t)] \] (6.232)

or
\[ y(t) = e^{-\gamma t} \left( A'' \cos (\Omega t - \gamma) \right). \] (6.233)

It shows under-damped oscillatory motion.

6.4.2 None homogeneous ODE’s with Constant Coefficients

An $n^{th}$-order none homogeneous ODE with constant coefficient has the form
\[ a_0 y + a_1 \frac{dy}{dx} + a_2 \frac{d^2y}{dx^2} + a_3 \frac{d^3y}{dx^3} \ldots a_{n-1} \frac{d^{n-1}y}{dx^{n-1}} + a_n \frac{d^n y}{dx^n} = f(x). \] (6.234)

In this section, we will see the procedure that is built upon on the methods we developed in the previous sections for HODE and an educated guess to part of the general solution related to the none homogeneity of the DE. This procedure consist of the following successive steps.

1-st Find the general solution to the HODE (i.e. with $f(x) = 0$). This solution is called the complementery function, or the homogeneous solution, denoted by $y_h(x)$. Note that this solution contains $n$ constants of integration for the $n^{th}$-order ODE.

2-nd Find in any way you can a particular solution, $y_p(x)$, that satisfies the full none homogeneous ODE. Note that this function contains no arbitrary constants of integration.

3-rd The general solution to the none homogeneous ODE is then given by
\[ y(x) = y_h(x) + y_p(x). \] (6.235)

4-th Any boundary conditions given to the problem must then be applied to the general solution $y(x)$. 
Example 13 Find the general solution to the following differential equation:

\[ \frac{d^2s}{dt^2} - 3 \frac{ds}{dt} + 2s = 16. \]  

(6.236)

Solution:

Step 1: Find the homogeneous solution. That means determine the solution for

\[ \frac{d^2s}{dt^2} - 3 \frac{ds}{dt} + 2s = 0. \]  

(6.237)

Noting that the indicial equation is

\[ k^2 - 3k + 2 = 0 \Rightarrow k_1 = 2, k_2 = 1 \]  

(6.238)

the homogeneous solution can be written as

\[ s_h = Ae^{2t} + Be^t. \]  

(6.239)

Step 2: By inspecting the none homogeneous ODE, I guess a particular solution of the form

\[ s_p = C \Rightarrow \frac{ds_p}{dt} = \frac{d^2s_p}{dt^2} = 0, \]  

(6.240)

so that when we substitute it back into the none homogeneous ODE in Eq. (6.236), we find

\[ 2C = 16 \Rightarrow C = 8. \]  

(6.241)

Thus the particular solution is found to be

\[ s_p = 8. \]  

(6.242)

Step 3: Write the general solution

\[ s(t) = s_h(t) + s_p(t) \Rightarrow s(t) = Ae^{2t} + Be^t + 8. \]  

(6.243)

Step 4: Apply the boundary conditions if there are any. Here we are not given any.

Example 14 Find the general solution to the following differential equation:

\[ \frac{d^2R}{dx^2} + 4R = \sin (3x) \]  

(6.244)

Solution:

Step 1: Find the homogeneous solution for

\[ \frac{d^2R}{dx^2} + 4R = 0. \]  

(6.245)
Noting that the roots to the indicial equation are
\[ k^2 + 4 = 0 \Rightarrow k_1 = 2i, k_2 = -2i \]
the homogeneous solution can be written as
\[ R_h(x) = A \cos(2x) + B \sin(2x). \]  

**Step 2:** By inspecting the given none homogeneous ODE, I made a guess (an educated guess) to the particular solution that has the form
\[ R_p = C \sin(3x). \]

*Note that this guess will be valid because \( \pm 3i \) is different from the roots of the indicial equation. See the next example.*

Since in the particular solution we are not allowed to have arbitrary constant of integration, we must determine \( C \). Upon substituting Eq. (6.248) into Eq. (6.244), this constant can be determined to be
\[
-9C \sin(3x) + 4C \sin(3x) = \sin(3x) \Rightarrow -5C \sin(3x) = \sin(3x)
\]
\[ \Rightarrow C = -\frac{1}{5}. \]

Thus the particular solution becomes
\[ H_p = -\frac{1}{5} \sin(3x). \]

**Step 3:** Write the general solution
\[ R(x) = R_h(x) + R_p(x) \Rightarrow R(x) = A \cos(2x) + B \sin(2x) - \frac{1}{5} \sin(3x). \]

**Step 4:** Apply the boundary conditions if there are any. Here there are no any.

**Example 16** Find the general solution to the following differential equation:
\[
\frac{d^2y}{dx^2} + 4y = \sin(2x)
\]

**Solution:**

**Step 1:** Find the solution for the homogeneous LDE
\[
\frac{d^2y}{dx^2} + 4y = 0.
\]

The roots to the indicial equation
\[ k^2 + 4 = 0 \]
are
\[ k_1 = 2i, \quad k_2 = -2i. \] (6.255)

Then the homogeneous part of the solution can be expressed as
\[ y_h(x) = A \cos(2x) + B \sin(2x), \] (6.256)
or preferably
\[ y_h(x) = A'e^{2ix} + B'e^{-2ix}. \] (6.257)

**Step 2:** When the non-homogenous part \( f(x) \) is a trigonometric or hyperbolic function (i.e. \( \sin(ax), \cos(ax), \sinh(ax), \) or \( \cosh(ax) \)), before we try to find the particular solution it is recommended to express the functions as exponential function applying Euler’s formula. Thus for function in Eq. (6.252), one should write
\[ f(x) = \sin(2x) = \frac{1}{2i} (e^{2ix} - e^{-2ix}). \] (6.258)

Now comparing Eqs. (6.257) and (6.258), we note that values to the exponents to the Homogeneous solution (the roots of the indicial equation) are the same as the exponents to the function, \( f(x) \). Under these circumstances, one can make an educated guess to the particular solution based on the property of an exponential function and the order of the differential equation. For example, for the DE in Eq. (6.252), if one guesses a particular solution
\[ y_p(x) = C_1 e^{2ix} + C_2 e^{-2ix} \]
which has the same form as the homogeneous solution in Eq. (6.257), we find
\[ \frac{d^2 y_p}{dx^2} + 4y_p = -4y_p + 4y_p = 0, \]
as one would have guessed. Therefore, we must guess a function that must have a different form the homogenous solution. In this case, our educated guess be
\[ y_p(x) = x \left( C_1 e^{2ix} + C_2 e^{-2ix} \right), \] (6.259)
that leads to
\[ \frac{dy_p}{dx} = C_1 e^{2ix} + C_2 e^{-2ix} + 2ix \left(C_1 e^{2ix} - C_2 e^{-2ix}\right) \]
\[ \Rightarrow \frac{d^2 y_p}{dx^2} = 2i \left(C_1 e^{2ix} - C_2 e^{-2ix}\right) + 2i \left(C_1 e^{2ix} - C_2 e^{-2ix}\right) \\
-4x \left(C_1 e^{2ix} + C_2 e^{-2ix}\right) \]
\[ \Rightarrow \frac{d^2 y_p}{dx^2} + 4y_p = 4i \left(C_1 e^{2ix} - C_2 e^{-2ix}\right). \] (6.260)

Substituting this result into Eq. (??), we find
\[ 4i \left(C_1 e^{2ix} - C_2 e^{-2ix}\right) = \frac{1}{2i} (e^{2ix} - e^{-2ix}). \]
There follows that
\[ C_1 = C_2 = \frac{-1}{8}. \] (6.262)

Therefore, the particular solution becomes
\[ y_p (x) = \frac{-1}{8} [xe^{2ix} + xe^{-2ix}], \] (6.263)

that can be put in the form
\[ y_p (x) = \frac{-x}{4} \left[ \frac{e^{2ix} + e^{-2ix}}{2} \right] = \frac{-x}{4} \cos (2x). \] (6.264)

**Step 3:** Write the general solution,
\[ y (x) = y_h (x) + y_p (x) \]
\[ y (x) = A \cos (2x) + B \sin (2x) - \frac{x}{4} \cos (2x) \] (6.265)

**Step 3:** In this example also we are not given boundary conditions.

### 6.5 The method of superposition and the particular solution.

In the previous example the particular solution in Eq. (6.259) that we came up with an educated guess that can be put in the form
\[ y_p (x) = y_{p_1} (x) + y_{p_2} (x), \] (6.266)

where
\[ y_{p_1} (x) = P_n (x) e^{2ix}, y_{p_2} (x) = Q_n (x) e^{-2ix}. \] (6.267)

with \( P_n (x) = C_1 x \) and \( Q_n (x) = C_2 x \) are polynomial functions with degree, \( n = 1 \). To determine the particular solution, we also wrote the function, \( f (x) = \sin (2x) \), as a sum of two functions such that
\[ \frac{d^2 y_p (x)}{dx^2} + 4y_p (x) = f (x) = f_1 (x) + f_2 (x), \] (6.268)

where
\[ f_1 (x) = P_m (x) e^{2ix}, f_2 (x) = Q_m (x) e^{-2ix}, \] (6.269)

with \( P_m (x) = \frac{1}{2} \) and \( Q_m (x) = -\frac{1}{2} \) are polynomial of degree, \( m = 0 \). In a similar way, the particular solutions to
\[ \frac{d^2 y_{p_1} (x)}{dx^2} + 4y_{p_1} (x) = f_1 (x) \quad \text{and} \quad \frac{d^2 y_{p_2} (x)}{dx^2} + 4y_{p_2} (x) = f_2 (x), \] (6.270)

can be guess to be
\[ y_{p_1} (x) = P'_n (x) e^{2ix}, y_{p_2} (x) = Q'_n (x) e^{-2ix}. \] (6.271)
with \( P'_n(x) = C'_1 x \) and \( Q'_n(x) = C'_2 x \) are polynomial functions with degree, \( n = 1 \), respectively. Upon adding these two equations, we find a none homogeneous ODE in Eq. (6.272) has the same form as Eq. (6.268)

\[
\frac{d^2y_p}{dx^2} + 4y_p = f_1(x) + f_2(x) = f(x),
\]

where

\[
y_p(x) = 2y_{p_1}(x) + 2y_{p_2}(x) = P''_n(x) e^{2ix} + Q''_n(x) e^{-2ix},
\]

with \( P''_n(x) = 2C'_1 x = C''_1 x \) and \( Q''_n(x) = 2C'_2 x = C''_2 x \) are still the same polynomials with the same degree, \( n = 1 \) obtained the previous example by directly considering the function \( f(x) \). This can be generalized for an n-th order none homogeneous ODE

\[
\sum_n a_n \frac{d^n y}{dx^n} = f(x)
\]

that can be written as

\[
\sum_n a_n \frac{d^n y}{dx^n} = f_1(x) + f_2(x).
\]

The particular solution can be determined the principle of superposition that states

\[
y(x) = y_h(x) + y_{p_1}(x) + y_{p_2}(x)
\]

where \( y_{p_1}(x) \) and \( y_{p_2}(x) \) satisfy the ODE with the non-homogeneous terms \( f_1(x) \) and \( f_2(x) \),

\[
\sum_n a_n \frac{d^n y}{dx^n} = f_1(x)
\]

and

\[
\sum_n a_n \frac{d^n y}{dx^n} = f_2(x),
\]

respectively.

From what we discussed in relation to the previous example, we can also state that for a second-order Non-homogenous second-order ODE that can be expressed as

\[
\left( \frac{d}{dx} - a \right) \left( \frac{d}{dx} - b \right) y(x) = f_1(x) = e^{cx} P_n(x),
\]

where \( P_n(x) \) is a polynomial of degree \( n \), the particular solution of such kind of non-homogenous LDE is given by

\[
y_p(x) = \begin{cases} 
  e^{cx} Q_n(x), & \text{if } c \neq a \& c \neq b \\
  xe^{cx} Q_n(x), & \text{if } c = a \text{ or } b \& a \neq b \\
  x^2 e^{cx} Q_n(x), & \text{if } c = a = b
\end{cases}
\]

(6.280)
6.5. THE METHOD OF SUPERPOSITION AND THE PARTICULAR SOLUTION

where
\[ Q_n(x) = a_0 + a_1 x + a_2 x^2 + ... a_n x^n. \]  

(6.281)

is a polynomial of the same degree as \( P_n(x) \) with undetermined coefficients determined by substituting the particular solution to the given non-homogenous DE.

**Example 18** Using the principle of superposition solve the following differential equation:

\[ \frac{d^2 y}{dx^2} - y = 2 \sin(x) + 4x \cos(x) \]  

(6.282)

**Solution:** For the homogenous DE

\[ \frac{d^2 y}{dx^2} - y = 0 \]  

(6.283)

the indicial equation

\[ k^2 - 1 = 0 \]  

(6.284)

gives

\[ k_1 = 1, k_2 = -1. \]  

(6.285)

The the solution for the homogenous DE can be written as

\[ y_h(x) = C_1 e^x + C_2 e^{-x}. \]  

(6.286)

For the particular solutions first we rewrite the DE

\[ \frac{d^2 y}{dx^2} - y = 2 \sin(x) + 4x \cos(x) \]  

(6.287)

as

\[ \frac{d^2 y}{dx^2} - y = 2 \left( \frac{e^{ix} - e^{-ix}}{2i} \right) + 4x \left( \frac{e^{ix} + e^{-ix}}{2} \right) \]  

(6.288)

\[ \left( \frac{d}{dx} + 1 \right) \left( \frac{d}{dx} - 1 \right) = (2x - i) e^{ix} + (2x + i) e^{-ix} \]  

(6.289)

or

\[ \left( \frac{d}{dx} + 1 \right) \left( \frac{d}{dx} - 1 \right) = f_1(x) + f_2(x), \]  

(6.290)

where

\[ f_1(x) = (2x - i) e^{ix}, \quad f_2(x) = (2x + i) e^{-ix}, \]  

(6.291)

and

\[ a = 1, b = -1 \]  

(6.292)

Then for the non-homogenous DE with the function \( f_1(x) \)

\[ \left( \frac{d}{dx} + 1 \right) \left( \frac{d}{dx} - 1 \right) = f_1(x) = (2x - i) e^{ix} \]  

(6.293)
noting that 
\[ c = i \neq a, b \] (6.294)
we may write the particular solution as
\[ y_{p_1}(x) = (A + Bx)e^{ix}. \] (6.295)
that gives
\[
\frac{dy_{p_1}}{dx} = Be^{ix} + i(A + Bx)e^{ix} = Be^{ix} + iy_{p_1}(x),
\] (6.296)
\[
\frac{d^2y_{p_1}}{dx^2} = \frac{dB}{dx}e^{ix} + i\frac{dy_{p_1}}{dx} = iBe^{ix} + i(Be^{ix} + iy_{p_1}(x))
= 2iBe^{ix} - y_{p_1}(x) \] (6.297)
Substituting these equations into
\[
\frac{d^2y_{p_1}(x)}{dx^2} - y_{p_1}(x) = f_1(x) = (2x - i)e^{ix},
\] (6.298)
we find
\[
2iBe^{ix} - 2y_{p_1}(x) = (2x - i)e^{ix} \Rightarrow 2iBe^{ix} - 2(A + Bx)e^{ix}
= (2x - i)e^{ix}
\] (6.299)
\[
\Rightarrow [-2Bx + 2Bi - 2A]e^{ix} = (2x - i)e^{ix}. \] (6.300)
There follows that
\[
-2B = 2, 2Bi - 2A = -i \Rightarrow B = -1 \Rightarrow -2i - 2A = -i
\Rightarrow B = -1, A = -\frac{i}{2} \] (6.301)
and the first particular solution can be written as
\[ y_{p_1}(x) = -\left(\frac{i}{2} + x\right)e^{ix}. \] (6.302)
Similarly, for the second function we may write the second particular solution
\[ y_{p_2}(x) = (A' + B'x)e^{-ix} \] (6.303)
which gives
\[
\Rightarrow \frac{dy_{p_2}}{dx} = B'e^{-ix} - iy_{p_2}(x) \Rightarrow \frac{d^2y_{p_2}}{dx^2} = -2iB'e^{-ix} - y_{p_2}(x). \] (6.304)
Then
\[
\frac{d^2y_{p_2}(x)}{dx^2} - y_{p_2}(x) = f_2(x). \] (6.305)
becomes
\[-2iB' e^{-ix} - 2(A' + B'x)e^{-ix} = [-2B'x - 2B'i - 2A] e^{ix}
\]
\[= (2x + i) e^{-ix} \quad (6.306)\]

There follows that
\[-2B' = 2, \quad -2Bi - 2A = i \Rightarrow B' = -1 \Rightarrow 2i - 2A = i\]
\[\Rightarrow B' = -1, A' = \frac{i}{2} \quad (6.307)\]

and the second particular solution is
\[y_{p2}(x) = \left(\frac{i}{2} - x\right) e^{-ix} \quad (6.308)\]

Therefore, the general solution is given by
\[y(x) = C_1 e^x + C_2 e^{-x} - \left(\frac{i}{2} + x\right) e^{ix} + \left(\frac{i}{2} - x\right) e^{-ix} \quad (6.309)\]
or
\[y(x) = C_1 e^x + C_2 e^{-x} + \sin(x) - 2x \cos(x). \quad (6.310)\]

Using Mathematica!

\[
\text{In}[5]:= \text{DSolve}[[y''[x] - y[x] =
2 \sin[x] + 4 x \cos[x], y, x] \\
\text{Out}[5]= \{\{y \rightarrow \text{Function} \{\{x\}, e^x C[1] +
\quad e^{-x} C[2] - 2 x \cos [x] + \sin [x] \}\}\}
\]

*IS THAT NOT MUCH EASIER?*

### 6.6 The method of successive integration

In the previous section the methods introduced to determine the general solutions whether for homogeneous or none homogeneous ODE with constant coefficients is based on an educated guess. In this section, we will derive a general an integral expression for the general solution to a second-order none homogeneous ODE by successively integrating the ODE. The general solutions determined by an educated guess in the previous sections will be shown that it can be derived from this general integral expression. To this end, we put the second-order ODE with constant coefficients
\[a_2 \frac{d^2 y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = F(x) \quad (6.311)\]
in the form
\[ \left( a_2 \frac{d^2}{dx^2} + a_1 \frac{d}{dx} + a_0 \right) y = \left( \frac{d}{dx} - a \right) \left( \frac{d}{dx} - b \right) y = F(x), \] (6.312)

where \( a \) and \( b \) are the roots for a quadratic equation given by
\[ a = \frac{-a_1 - \sqrt{a_1^2 - 4a_2a_0}}{a_2}, \quad b = \frac{-a_1 + \sqrt{a_1^2 - 4a_2a_0}}{a_2}. \] (6.313)

Introducing a new variable \( u(x) \) defined by
\[ u(x) = \left( \frac{d}{dx} - b \right) y(x), \] (6.314)

Eq. (6.312) can be rewrite as
\[ \left( \frac{d}{dx} - a \right) u(x) = \frac{du}{dx} - au = F(x). \] (6.315)

Multiplying this equation by \( e^{-ax} \), we have
\[ e^{-ax} \frac{du}{dx} - ae^{-ax} u = e^{-ax} F(x), \] (6.316)

so that applying the product rule, one can write
\[ \frac{d}{dx} \left[ u e^{-ax} \right] = e^{-ax} F(x). \] (6.317)

Integrating this equation with respect to \( x \), we find
\[ u e^{-ax} = \int e^{-ax'} F(x') \, dx' + C_1 \]
\[ \Rightarrow u = e^{ax} \int e^{-ax'} F(x') \, dx' + C_1 e^{ax}, \] (6.318)
that leads to
\[ u(x) = e^{ax} \int e^{-ax'} F(x') \, dx' + C_1 e^{ax}, \] (6.319)

where \( C_1 \) is a constant of integration. Now substituting Eq. (6.319) into Eq. (6.314), one can write
\[ \frac{dy}{dx} - by = e^{ax} \int e^{-ax'} F(x') \, dx' + C_1 e^{ax}, \] (6.320)

and multiplying by \( e^{-bx} \), we find
\[ e^{-bx} \frac{dy}{dx} - be^{-bx} y = e^{(a-b)x} e^{ax} \int e^{-ax'} F(x') \, dx' + C_1 e^{(a-b)x}. \] (6.321)
Once again applying the product rule this can be put in the form
\[
\frac{d}{dx} (e^{-bx} y) = e^{(a-b)x} \int e^{-ax'} F(x') \, dx' + C_1 e^{(a-b)x}, \tag{6.322}
\]
so that integrating with respect to \( x \), we can write
\[
e^{-bx} y = \int e^{(a-b)x''} \int e^{-ax'} F(x') \, dx' \, dx'' + C_1 \int e^{(a-b)x''} \, dx'' + C_2, \tag{6.323}
\]
where again \( C_2 \) is a constant of integration. Thus the solution to the DE in Eq. (6.312) can be expressed as
\[
y(x) = e^{bx} \int e^{(a-b)x''} \int e^{-ax'} F(x') \, dx' \, dx'' + C_1 e^{bx} \int e^{(a-b)x''} \, dx'' + C_2 e^{bx}, \tag{6.324}
\]
where \( a \) and \( b \) are given by Eq. (6.313). Noting that the first term involving the function \( F(x) \) is a result of the none homogeneity of the DE and the second and third terms are independent of this function, we can rewrite Eq. (6.324) as
\[
y(x) = y_H(x) + y_p(x), \tag{6.325}
\]
where
\[
y_H(x) = C_1 e^{bx} \int e^{(a-b)x''} \, dx'' + C_2 e^{bx}, \tag{6.326}
\]
and
\[
y_p(x) = e^{bx} \int e^{(a-b)x''} \int e^{-ax'} F(x') \, dx' \, dx'' \tag{6.327}
\]
are the homogeneous and particular solutions, respectively. Next we will use Eqs. (??) and (??) to derive the solutions we determined in the previous sections using an educated guess.

**Problem 1:** Derive the integral expression for 3-rd and 4-th order none homogeneous ODE. From the results obtained try to write an expression for an n-th order none homogeneous ODE with constant coefficients.

**Problem 2:** Applying the general solution derived for a first-order ODE (Eq. (6.128) in section 6.3, derive Eqs. (6.319) and (6.324).

**A. HODE with none degenerate roots:** We recall that for a second-order homogeneous ODE \((F(x) = 0)\)
\[
a_2 \frac{d^2 y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = 0, \tag{6.328}
\]
that leads to an indicial equation
\[ a_2 k^2 + a_1 k + a_0 = 0 \]  
(6.329)

the roots are given by Eq. (6.313). When these roots are none degenerate \((a \neq b)\), we saw that the general solution to DE is given by
\[ y(x) = C_1 e^{ax} + C_2 e^{bx}. \]  
(6.330)

The result in Eq. (6.324), for \(F(x_0) = 0\); becomes
\[ y(x) = C_1 \int_0^x e^{(a-b)x''} dx'' + C_2 e^{bx}, \]  
(6.331)

and upon carrying out the integration, one finds
\[ y(x) = \frac{C_1 e^{bx} e^{(a-b)x}}{a - b} + C_2 e^{bx} = \frac{C_1 e^{ax}}{a - b} + C_2 e^{bx} \]
\[ = C_1 e^{ax} + C_2 e^{bx}. \]  
(6.332)

where we have absorbed the constant \(1/(a - b)\) in \(C_1\). This is the same as Eq. (6.330).

B. HODE with degenerate roots: For a homogeneous ODE that leads to an indicial equation with degenerate roots \((i.e. a = b)\), Eq. (6.331) becomes
\[ y(x) = C_1 e^{bx} \int_0^x dx'' + C_2 e^{bx} = (C_2 + C_1 x) e^{bx}, \]  
(6.333)

which is in agreement with Eq. (6.199).

C. None HODE: We have seen that the particular solutions, \(y_p(x)\), for the second-order none HODE
\[ a_2 \frac{d^2 y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = \left( \frac{d}{dx} - a \right) \left( \frac{d}{dx} - b \right) e^{cx} P_n(x), \]  
(6.334)

with a polynomial \(P_n(x)\) of degree \(n\), were determined based on an educated guess. For such none HODE, depending on the degeneracy of the values for \(a, b,\) and \(c\), we saw that the particular solutions are given by
\[ y_p(x) = \begin{cases} 
  e^{cx} Q_n(x), & \text{when } c \neq a \text{ & } c \neq b, \\
  xe^{cx} Q_n(x), & \text{when } c = a \text{ or } b \text{ & } a \neq b, \\
  x^2 e^{cx} Q_n(x), & \text{when } c = a = b.
\end{cases} \]  
(6.335)

Next we will derive these particular solutions from Eq. (6.327). To this end, using the expression for n-th degree polynomial function, \(P_n(x)\),
\[ P_n(x) = \sum_{n=0}^n A_i x^i, \]  
(6.336)
one can write the function, \( F(x) \), that makes the ODE none homogeneous, as

\[
F(x) = e^{cx} P_n(x) = e^{cx} \sum_{n=0}^{n} A_i x^n. \tag{6.337}
\]

Then the particular solution in Eq. (6.327) can be expressed as,

\[
y_p(x) = e^{bx} \int x \left[ e^{(a-b)x'} \sum_{n=0}^{n} A_i \int x'' \ e^{(c-a)x} x^n \ dx' \right] dx''. \tag{6.338}
\]

where

\[
\alpha = c - a \quad \text{and} \quad \beta = a - b. \tag{6.339}
\]

Next we will evaluate Eq. (6.338) to derive the particular solutions given in Eq. (6.335).

(i.) When \( c \neq a \) \& \( c \neq b \): Under this condition, \( \alpha = c - a \neq 0 \). Under this condition one can use the relation

\[
\frac{d^i}{d\alpha^i} [e^{\alpha x}] = e^{\alpha x} x'^i, \tag{6.340}
\]

and show that

\[
\int x'' e^{\alpha x} x'^i \ dx' = \frac{d^i}{d\alpha^i} \int x'' e^{\alpha x} \ dx' = \frac{d^i}{d\alpha^i} \left( e^{\alpha x''} \right). \tag{6.341}
\]

Substituting this equation into Eq. (6.338), we can write the particular solution as

\[
y_p(x) = e^{bx} \sum_{i=0}^{n} A_i \int x \left[ e^{\beta x''} \frac{d^i}{d\alpha^i} \left( e^{\alpha x''} \right) \right] dx''. \tag{6.342}
\]

Upon carrying out the integration, we find

\[
y_p(x) = e^{bx} \sum_{i=0}^{n} A_i \frac{d^i}{d\alpha^i} \left[ \frac{e^{(\alpha + \beta) x}}{\alpha + \beta} \right]. \tag{6.343}
\]
Noting that
\[
\frac{d}{d\alpha} \left[ \frac{e^{(\alpha+\beta)x}}{\alpha+\beta} \right] = \frac{e^{(\alpha+\beta)x}}{\alpha+\beta} = C_0 e^{(\alpha+\beta)x},
\]
\[
\frac{d}{d\alpha} \left[ \frac{e^{(\alpha+\beta)x}}{\alpha+\beta} \right] = \frac{d}{d\alpha} \left[ -\frac{1}{(\alpha+\beta)^2} + \frac{x}{\alpha+\beta} \right] e^{(\alpha+\beta)x} = (C_0 + C_1 x) e^{(\alpha+\beta)x},
\]
\[
\frac{d^2}{d\alpha^2} \left[ \frac{e^{(\alpha+\beta)x}}{\alpha+\beta} \right] = \frac{d}{d\alpha} \left[ -\frac{1}{(\alpha+\beta)^2} + \frac{x}{\alpha+\beta} \right] e^{(\alpha+\beta)x} = \left( C_0 + C_1 x + C_2 x^2 \right) e^{(\alpha+\beta)x},
\]
(6.344)

one can easily show that
\[
\sum_{i=0}^{n} A_i \frac{d^i}{dx^i} \int e^{(\alpha+\beta)x''} dx'' = (C_0 + C_1 x + C_2 x^2 + \ldots + C_n x^n) e^{(c-a)x} = Q_n(x) e^{(c-a)x},
\]
(6.345)

where \(Q_n(x)\) is a polynomial of degree \(n\). Note that the constants \(A_i\) are absorbed in the constants \(C_i\). Thus, substituting, \(\alpha = c-a\) and \(\beta = a-b\) in Eq. (6.345), the particular solution in Eq. (6.343) becomes
\[
y_p(x) = e^{bx} Q_n(x) e^{(c-a-a-b)x} = Q_n(x) e^{cx}.
\]
(6.346)

(ii) When \(c = a\) \& \(c \neq b\): for this case we have \(\alpha = c-a = 0\) and Eq. (6.338) becomes
\[
y_p(x) = \sum_{n=0}^{n} A_i e^{bx} \int e^{(\alpha+\beta)x''} dx'' = \left( C_0 + C_1 x + C_2 x^2 + \ldots C_n x^n \right) e^{(c-a-a-b)x} = Q_n(x) e^{cx}.
\]
(6.347)

Substituting this into Eq. (6.327), the particular solution can be expressed as
\[
y_p(x) = \sum_{i=0}^{n} A_i \frac{d^i}{dx^i} \int e^{(\alpha+\beta)x''} dx'' \int e^{\beta x''} (x^{(i+1)}) dx'' = e^{bx} \sum_{j=1}^{n+1} \frac{A_{j-1}}{j} \int e^{(\alpha+\beta)x''} (x^{(j)}) dx''
\]
(6.348)

where we have introduced the summation index defined by \(j = i+1\).

Applying the relation in Eq. (6.340), one can easily show that
\[
y_p(x) = e^{bx} \sum_{j=1}^{n+1} \frac{A_{j-1}}{j} \frac{d^j}{d\beta^j} \int e^{(\alpha+\beta)x''} dx'' = e^{bx} \sum_{j=1}^{n+1} \frac{A_{j-1}}{j} \frac{d^j}{d\beta^j} \left[ \frac{e^{\beta x}}{\beta} \right].
\]
(6.349)
Following a similar procedure we used in (ii) earlier, one can show that
\[\sum_{j=1}^{n+1} A_{j-1} \frac{d^{j}}{d\beta^{j}} \left[ \frac{e^{\beta x}}{\beta} \right] = (C_{1}x + C_{2}x^{2} + ...C_{n+1}x^{n+1}) e^{\beta x}\]
\[= x (C_{0} + C_{1}x + C_{2}x^{2} ...C_{n}x^{n}) e^{\beta x}.\] (6.350)

Thus the particular solution in Eq. (6.349) becomes
\[y_{p}(x) = e^{bx}xe^{\beta x}Q_{n}(x) = xe^{cx}Q_{n}(x),\] (6.351)
where \(Q_{n}(x)\) is a polynomial of degree \(n\). This result is the same as the solution in Eq. (6.335) when \(c = a \quad \& \quad c \neq b\) or \(c = b \quad \& \quad c \neq a\).

(iii.) When \(c = a = b\) : for this case we have \(\alpha = c-a = 0\) and \(\beta = a - b\) and Eq. (6.348) becomes
\[y_{p}(x) = e^{bx} \sum_{j=1}^{n+1} A_{j-1} \int_{x}^{x} (x^{n})^{j} dx^{n}.\] (6.352)

Upon integrating this, we find
\[y_{p}(x) = e^{bx} \sum_{j=1}^{n+1} A_{j-1} \frac{(x^{n})^{j+1}}{j(j+1)},\] (6.353)
that can be put in the form
\[y_{p}(x) = (C_{2}x^{2} + C_{2}x^{3} + ...C_{n}x^{n+2}) e^{bx} = x^{2}e^{cx}Q_{n}(x),\] (6.354)
where \(Q_{n}(x)\) is a polynomial of degree \(n\). This result is also the same as the solution in Eq. (6.335) when \(c = a = b\).

**Example 17** A forced Harmonic oscillator: Consider an ideal spring with spring constant \(k = 8N/m\) with a mass \(m = 2kg\) attached to its one end with the other end fixed. The spring and mass are sitting on a frictionless surface. The mass is constantly acted by a periodic external force given by, \(F_{e} = 2 \sin (3t) \hat{x}\). What would be the amplitude of the oscillation. Initially the mass is at rest with the spring at its equilibrium position. (i.e. the spring is neither stretched nor compressed.)
Solution: In order to find the amplitude of oscillation one needs to solve the equation of motion of the mass. Applying Newton’s second law one can write the equation of motion for the mass, \( m \), as

\[
m \frac{d^2 x}{dt^2} = F_e - kx = 2 \sin (3t) - kx. \tag{6.355}
\]

Using the values for the spring constant, \( k = 8 \text{N/m} \), and mass, \( m = 2 \text{kg} \), that can be put in the form

\[
d^2 x + 4x = \left( \frac{d^2}{dt^2} + 4 \right) x = \left( \frac{d}{dt} + 2i \right) \left( \frac{d}{dt} - 2i \right) x = \sin (3t). \tag{6.356}
\]

In view of the general solutions in Eqs. (6.326) and (6.327) that we derived by successively integrating the second-order none homogeneous ODE given in Eq. (6.312), the homogeneous and the particular solutions to Eq. (6.356) can be expresses as

\[
x_H (t) = C_1 e^{bt} \int^t e^{(a-b)t''} dt'' + C_2 e^{bt} \tag{6.357}
\]

and

\[
x_p (t) = e^{bt} \int^t \left[ e^{(a-b)t''} \int^t e^{-at'} F (t') dt' \right] dt''. \tag{6.358}
\]

From Eq. (6.356), we have

\[
a = 2i, b = -2i, \text{ and } F (t) = \sin (3t) = \frac{e^{3it} - e^{-3it}}{2i}, \tag{6.359}
\]

so that Eqs. (6.357) becomes

\[
x_H (t) = C_1 e^{-2it} \int^t e^{4it''} dt'' + C_2 e^{-2it} = \frac{C_1}{4i} e^{2it} + C_2 e^{-2it}. \tag{6.360}
\]

Using Euler’s formula the homogeneous solutions can be expressed as

\[
x_H (t) = \frac{C_1}{4i} [\cos (2t) + 2i \sin (2t)] + C_2 [\cos (2t) - 2i \sin (2t)]
\]

\[
= \left( \frac{C_1}{4i} + 2iC_2 \right) \cos (2t) + \left( \frac{C_1}{4i} - 2iC_2 \right) \sin (2t), \tag{6.361}
\]

we can rewrite as this equation as

\[
x_H (t) = A \cos (2t) + B \sin (2t). \tag{6.362}
\]

Substituting the values for \( a, b, \) and \( F (t) \) from Eq. (6.359) into Eq. (6.358), the particular solution, becomes

\[
x_p (t) = e^{-2it} \int^t e^{4it''} \left[ e^{2it'} \left( \frac{e^{3it'} - e^{-3it'}}{2i} \right) \right] dt''
\]

\[
= \frac{e^{-2it}}{2i} \int^t e^{4it''} \left[ e^{it'} - e^{5it'} \right] dt''. \tag{6.363}
\]
Upon carrying out the inner integral, we have
\[ x_p(t) = \frac{e^{-2it}}{2i} \int e^{4it''} \left( \frac{e^{i5t''}}{i} + \frac{e^{-i5t''}}{5i} \right) dt'' = \frac{e^{-2it}}{2i} \int \left( \frac{e^{i5t''}}{i} + \frac{e^{-i5t''}}{5i} \right) dt'', \] (6.364)
so that one can also carry out similar integration to this equation and find
\[ x_p(t) = \frac{e^{-2it}}{2i} \left( \frac{e^{5it}}{5i^2} - \frac{e^{-it}}{5i^2} \right) = \frac{1}{5} \left( \frac{e^{3it} - e^{3it}}{2i} \right) = -\frac{1}{5} \sin (3t). \] (6.365)
Using the results in Eqs. (6.362) and (6.365), the general solution, \( x(t) = x_H(t) + x_p(t) \), becomes
\[ x(t) = A \cos (2t) + B \sin (2t) - \frac{1}{5} \sin (3t). \] (6.366)
Initially \( t = 0 \), the mass is at rest with the spring at its equilibrium position. (i.e. the spring is neither stretched nor compressed.). This means for the position of the mass given by Eq. (6.366), we find
\[ x(t = 0) = 0 \implies A = 0, \] (6.367)
and for the velocity, that can easily be determined from \( x(t) \),
\[ v(t) = \frac{dx(t)}{dt} = 2 (-A \sin (2t) + B \cos (2t)) - \frac{3}{5} \cos (3t), \] (6.368)
we also obtain
\[ v(t = 0) = 0 \implies 2B - \frac{3}{5} = 0 \implies B = \frac{3}{10}. \] (6.369)
Therefore, with the given initial conditions, the position is given by
\[ x(t) = \frac{3}{10} \sin (2t) - \frac{1}{5} \sin (3t). \] (6.370)
and the velocity
\[ v(t) = \frac{1}{5} [2 \cos (2t) - 3 \cos (3t)]. \] (6.371)
Using Mathematica:
6.7 Partial Differential Equations

In the previous sections we saw that physical properties of some real systems can be described by homogeneous or non-homogeneous ordinary differential equations (ODE) of unknown functions of one variable. The techniques for solving such kind of differential equation were also discussed in these sections. Generally, most real physical systems or process are described by unknown functions of two or more variables that must be determined from a homogeneous or non-homogeneous Partial Differential Equation (PDE).

A Partial Differential Equation (PDE) is a differential equation involving partial derivatives of an unknown function with respect to two or more independent variables. There are various physical processes that requires use of partial differential equations. Some are listed below:

1. The gravitation potential of a given mass distribution described by given mass density or electrical potential of a given charge distribution with a given charge density satisfy the partial differential equation (for example in Cartesian coordinates) given by

\[
\frac{\partial^2 V(x, y, z)}{\partial x^2} + \frac{\partial^2 V(x, y, z)}{\partial y^2} + \frac{\partial^2 V(x, y, z)}{\partial z^2} = \rho(x, y, z),
\]

where \(V(x, y, z)\) is the gravitation potential (electrical potential) of a given mass (charge) distribution and \(\rho(x, y, z)\) is the mass (charge) density. If you know the density, then by solving this partial differential equation you can determine the potential at some point in space. This PDE is known as the Poisson’s equation. If \(\rho(x, y, z) = 0\), it is called Laplace’s equation.
2. The Schrödinger equation:

\[
-\frac{\hbar^2}{2m} \left[ \frac{\partial^2 \Psi(x,y,z,t)}{\partial x^2} + \frac{\partial^2 \Psi(x,y,z,t)}{\partial y^2} + \frac{\partial^2 \Psi(x,y,z,t)}{\partial z^2} \right] + U(x,y,z) \Psi(x,y,z,t) = -i\hbar \frac{\partial}{\partial t} \Psi(x,y,z,t),
\]

where $\Psi(x,y,z,t)$ is the wave function and $U(x,y,z)$ is the potential energy.

3. The wave equation:

\[
\frac{\partial^2 u(x,y,z,t)}{\partial x^2} + \frac{\partial^2 u(x,y,z,t)}{\partial y^2} + \frac{\partial^2 u(x,y,z,t)}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 u(x,y,z,t)}{\partial t^2},
\]

In this last section of this chapter, we will be introduced to the technique for solving PDE and find the unknown function as it applied to real physical systems or processes. This technique is known as the separation of variables technique described by the following successive steps that build upon the methods for solve ODE that we were introduced to in the previous sections.

1. Assume a product form for the solution.

2. Substitute the product function into the original PDE and simplify.

3. Divide both sides of the resulting DE by the product function.

4. Try to separate the variables, getting only one variable on one side of the equation.

5. The two sides of the equation must then be equal to the same constant, called the separation constant. (Think about what to name this constant—you can save yourself a lot of needless work by choosing a good name!)

6. Solve the resulting ordinary DE’s using techniques that you were introduced to in the previous sections.

7. The general solution is then a linear combination of all possible product-form solutions to the PDE, including all possible values of the separation constant!

8. Apply any BC’s to the general solution.

**Example 18 Traveling Waves on a Stretched String:**

(a) Show that the string supports a wave motion by deriving the wave equation.

\[
\frac{\partial^2 u(z,t)}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 u(z,t)}{\partial t^2},
\]
where

\[ v = \sqrt{\frac{T}{\mu}}, \]  

(6.376)

\( T \) is the tension on the string, and \( \mu \) is the linear mass density.

(b) Solve the 1-D traveling wave equation for a string of length \( L \) having both ends fixed. Which means, apply the BC’s

\[ u(y, t = 0) = u(y = 0, t) = u(y = L, t) = 0. \]  

(6.377)

Solution:

(a) *Because of the weight of the mass hanging* the string experiences a tension force \( T \). When is displaced from the equilibrium by the vibrator as shown in the figure below: If the string is displaced from the equilibrium, the net
transverse force on the segment between \( y \) and \( y + \Delta y \) can be expressed as
\[
\Delta F = T \left( \sin \left( \theta' \right) - \sin \left( \theta \right) \right)
\]
(6.378)

If we assume the displacement is small, we can replace \( \sin \) by \( \tan \),
\[
\Delta F = T \left( \tan \left( \theta' \right) - \tan \left( \theta \right) \right).
\]
(6.379)

Noting that
\[
\tan \left( \theta' \right) = \frac{\partial u}{\partial y} \bigg|_{y=y+\Delta y}, \tan \left( \theta \right) = \frac{\partial u}{\partial y} \bigg|_{y=y},
\]
(6.380)
we may write
\[
\Delta F = T \left( \frac{\partial u}{\partial y} \bigg|_{y=y+\Delta y} - \frac{\partial u}{\partial y} \bigg|_{y=y} \right).
\]
(6.381)

We recall that the definition of the first derivative of a function \( u(y) \) at \( y \)
\[
\frac{\partial u}{\partial y} \bigg|_{y=y} = \lim_{\Delta y \rightarrow 0} \left[ \frac{u(y + \Delta y) - u(y)}{\Delta y} \right]
\]
(6.382)
and the second derivative
\[
\frac{\partial^2 u}{\partial y^2} \bigg|_{y=y} = \lim_{\Delta y \rightarrow 0} \left[ \frac{\frac{\partial u}{\partial y} \bigg|_{y=y+\Delta y} - \frac{\partial u}{\partial y} \bigg|_{y=y}}{\Delta y} \right]
\]
(6.383)
\[
\Rightarrow \frac{\partial^2 u}{\partial y^2} \bigg|_{y=y} \Delta y = \lim_{\Delta y \rightarrow 0} \left[ \frac{\partial u}{\partial y} \bigg|_{y=y+\Delta y} - \frac{\partial u}{\partial y} \bigg|_{y} \right]
\]
(6.384)
so that for small \( \Delta y \), we can write
\[
\Delta F \simeq T \frac{\partial^2 u}{\partial y^2} \Delta y
\]
(6.385)

Noting that for the transverse motion, Newton’s second law can be written as
\[
\Delta F = ma = m \frac{\partial^2 u}{\partial t^2}
\]
(6.386)
If the string has a linear mass density \( \mu \), we can write \( m = \mu \Delta y \) so that
\[
\Delta F = \mu \Delta y \frac{\partial^2 u}{\partial t^2}
\]
(6.387)
Combining the two expressions for the force \( \Delta F \), we find
\[
T \frac{\partial^2 u}{\partial y^2} \Delta y = \mu \Delta y \frac{\partial^2 u}{\partial t^2}
\]
(6.388)
Which leads to
\[ \frac{\partial^2 u}{\partial y^2} = \frac{\mu \partial^2 u}{T \partial t^2} \]
or
\[ \frac{\partial^2 u}{\partial y^2} = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} \] (6.389)

where
\[ v = \sqrt{\frac{T}{\mu}}. \] (6.390)

Equation (6.389) is the one-dimensional wave equation. It describes a wave traveling in the y-direction with a speed, \( v \), given by Eq. (6.390) that depends on the magnitude of the tension on the string, \( T \), and the linear mass density of the string, \( \mu \).

(b) Let’s assume the solution of the PDE can be expressed as
\[ u(y, t) = Y(y) T(t) \] (6.391)
so that we can write
\[ T(t) \frac{d^2 Y(y)}{dy^2} = \frac{Y(y)}{v^2} \frac{d^2 T(t)}{dt^2} \] (6.392)

dividing both sides by \( Y(y) T(t) \)
\[ \frac{1}{Y(y)} \frac{d^2 Y(y)}{dy^2} = \frac{1}{v^2 T(t)} \frac{d^2 T(t)}{dt^2} = -k^2 \Rightarrow \frac{d^2 Y(y)}{dy^2} + k^2 Y(y) = 0, \]
\[ \frac{d^2 T(t)}{dt^2} + k^2 v^2 T(t) = 0 \] (6.393)

The solutions to these HDEs is given by
\[ Y(y) = A \cos(ky) + B \sin(ky), \quad T(t) = C \cos(kvt) + D \sin(kvt) \] (6.394)

Therefore the general solution can be expressed as
\[ u(y, t) = \sum_k [A_k \cos(ky) + B_k \sin(ky)] [C_k \cos(kvt) + D_k \sin(kvt)] \] (6.395)

Now we apply the boundary conditions. For \( u(y = 0, t) = 0 \), we have
\[ u(0, t) = \sum_k A_k (C_k \cos(kvt) + D_k \sin(kvt)) = 0 \] (6.396)

we have
\[ u(y, t) = \sum_k A_k (C_k \cos(kvt) + D_k \sin(kvt)) = 0 \] (6.397)

This can be zero for all time \( t \), if an only if when \( A_k = 0 \). This leads to simplified expression for the general solution
\[ u(y, t) = \sum_k B_k \sin(ky) [C_k \cos(kvt) + D_k \sin(kvt)]. \] (6.398)
Then applying the second boundary conditions for \( u(y, t = 0) = 0 \), we find
\[
  u(y, t = 0) = \sum_k (B_k \sin (ky)) C_k = 0 \Rightarrow C_k = 0. \tag{6.399}
\]
Thus the further simplified solution can be rewritten as
\[
  u(y, t) = \sum_k B_k D_k \sin (ky) \sin (kvt) = \sum_k B'_k \sin (ky) \sin (kvt) \tag{6.400}
\]
Then, applying the second boundary condition \( u(y = L, t) = 0 \), we have
\[
  u(L, t) = \sum_k B'_k \sin (kL) \sin (kvt) = 0 \Rightarrow \sin kL = 0
\]
\[
  \Rightarrow kL = n\pi \Rightarrow k = \frac{n\pi}{L}, \text{ where } n = 0, 1, 2, \ldots. \tag{6.401}
\]
Therefore, the general solution becomes
\[
  u(y, t) = \sum_{n=0} B'_n \sin (k_n y) \sin (k_n vt). \tag{6.402}
\]
where
\[
  k_n = \frac{n\pi}{L}, n = 0, 1, 2, 3, \ldots. \tag{6.403}
\]
This concludes the chapter to the introduction to differential equations.

### 6.8 Homework Assignment 12

1. Boas 8.1 #6
2. Boas 8.1 #7
3. Boas 8.2 #6
4. Boas 8.2 #32

### 6.9 Homework Assignment 13

1. Consider the first-order differential
\[
  2x \frac{dy}{dx} + y = 2x^{5/2}
\]
(a) Show that the differential equation is an exact differential equation or not.
(b) Write the differential equation in the form
\[
  \frac{dy}{dx} + P(x) y = Q(x)
\]
and determine the solution using the relation derived for such differential equation.
(c) Verify your solution using computer programing (Mathematica is recommended).

2. Do part (a)-(c) in problem 1 for the differential equation
\[ \frac{dy}{dx} + y \tanh x = 2e^x \]

3. Consider 10^9 Gallons volume salty lake. Suppose a salty water with a salt density
\[ \rho_{sw} = 5 \times 10^3 \text{lb/gal} \]

begins to flow at a rate of
\[ \frac{dV_{in}}{dt} = 4 \times 10^5 \text{gal/h.} \]

into the salty lake. At the same time the salty water in the lake begin to flow out at a rate of
\[ \frac{dV_{out}}{dt} = 10^5 \text{gal/h.} \]

(a) Determine the differential equation describing the volume of the salty water, \( V(t) \), in the lake. Solve the differential equation to find the volume of the salty water as a function of time.

(b) Find the differential equation that describes the mass, \( m(t) \) of the salt in the salty lake. Solve the differential equation to find the mass of the salt in the lake as function of time if the mass of the salt in the lake at the initial time, before the in and out flows begin, is \( m(0) = 10^7 \text{lb} \).

### 6.10 Homework Assignment 14

1. Find the general solution for the None-homogeneous DE
\[ \frac{d^2y}{dx^2} + \frac{dy}{dx} - 2y = e^{2x}. \]

2. Find the general solution for the None-homogeneous DE
\[ 5 \frac{d^2y}{dx^2} + 6 \frac{dy}{dx} + 2y = x^2 + 6x. \]

3. Find the general solution for the None-homogeneous DE
\[ \left( \frac{d^2}{dx^2} + 1 \right) y = 8x \sin x. \]

4. Using the method of superposition find the general solution to the DE
\[ \frac{d^2y}{dx^2} - 5 \frac{dy}{dx} + 6y = 2e^x + 6x - 5. \]
6.11 Homework Assignment 15

1. In classical physics, the motion of a particle can be described for all time once the “initial conditions” (position, velocity) of the particle’s motion are known. The so-called “equation of motion” in this case is Newton’s second law. In order to apply Newton’s second law, one must be able to specify all of the forces acting on the particle in question. In quantum mechanics, the equation of motion is called the Schrödinger equation. This equation is a partial differential equation involving a function called the wave function for the particle, denoted \( \Psi(x,t) \) (for 1-D problems), which contains all observable information about the particle’s motion (position, velocity, momentum, energy,...). Instead of precisely specifying all future motion of the particle, in quantum mechanics the quantity \( |\Psi(x,t)|^2 \) is used to predict the probability of a particular future motion. The 1-D Schrödinger equation is given by

\[
\frac{-\hbar^2}{2m} \frac{\partial^2 \Psi(x,t)}{\partial x^2} + U(x) \Psi(x,t) = i\hbar \frac{\partial \Psi(x,t)}{\partial t}
\]

where \( m \) is the mass of the particle, \( \hbar \) is Planck’s constant \( (1.05 \times 10^{-34} \text{Js}) \), \( i \) is the pure imaginary number, and \( U(x) \) is the net potential energy function of the particle at position \( x \). (Note that, since \( F_x = -dU(x)/dx \), specifying the potential energy function in quantum mechanics is equivalent to specifying all of the forces acting on the particle in classical mechanics.) The form of the Schrödinger equation given above is called the time-dependent (1-D) Schrödinger equation (TDSE) since it involves the full time-dependent wave function, \( \Psi(x,t) \).

(a) Using the separation of variables technique, and letting the separation constant be denoted \( E \) (which turns out to be the total energy of the particle), show that the resulting differential equation that is independent of time—the so-called time-independent Schrödinger equation (TISE)—has the (1-D) form

\[
-\frac{\hbar^2}{2m} \frac{d^2 \psi(x,t)}{dx^2} + U(x) \psi(x,t) = E \psi(x,t)
\]

where we have set \( \Psi(x,t) = \psi(x) T(t) \). In applications, the potential energy function \( U(x) \) must be specified, and the resulting TISE solved for the spatial wave function \( \psi(x) \).

(b) Also, show that the total wave function in this case has the form

\[
\Psi(x,t) = \psi(x) e^{-iEt/\hbar}.
\]
Chapter 7

Integral calculus-scalar functions

In this chapter we will be introduced to single and multiple integrals involving scalar functions. We focus on the methods for carrying out such integrals. In almost all branches of physics, integrations of scalar functions are often used to determine various quantities describing the static or dynamic properties of an object. The object could be in a microscopic particle like an electron in an atom or massive objects like a planet in the solar system. The simplicity of carrying out integrations of scalar functions depends on the geometry of the object and system that object is existing in. The methods of integrations in Cartesian coordinates for rectangular geometry and the methods of integrations in curvilinear coordinates for curved geometry can greatly simplify the integrations. We will be introduced the methods of integrations of single and multiple integrals in Cartesian and curvilinear coordinates and its applications in real physical problems.

7.1 Integration in Cartesian coordinates

We are interested in the methods of integrations of functions of single, two, and three variables in Cartesian coordinates.

7.1.1 Single and multiple integrals

A. Single integral and arc length

Consider the two points $P_1$ and $P_2$ in the x-y plane shown in Fig. 7.1.1 described by the Cartesian coordinates $(x_1, y_1(x_1))$ and $(x_2, y_2(x_2))$, respectively. You took two paths to get from the first to the second point. The first is a straight line path shown in red and the second is a curved path shown by the black curve. The slope
m(x) = \lim_{\Delta x \to 0} \frac{\Delta y(x)}{\Delta x} = \frac{dy(x)}{dx}. \quad (7.1)

is a constant for all x from the first to the second point along the first path but not along the second path. For any of these paths the net distance traveled along the y-direction, that can be determined from the line integral

\[ \Delta y = y_2(x_2) - y_1(x_1) = \int_{y_1(x_1)}^{y_2(x_2)} dy = \int_{x_1}^{x_2} m(x) \, dx. \quad (7.2) \]

are the same. Similarly, along the x-direction, the net distance given by

\[ \Delta x = x_2 - x_1 = \int_{y_1}^{y_2} m'(y) \, dy, \quad (7.3) \]

where

\[ m'(y) = \lim_{\Delta y \to 0} \frac{\Delta x(y)}{\Delta y} = \frac{dx(y)}{dy}. \quad (7.4) \]

are the same for the two paths.

However, for the total distances (arc length), l, for these two paths are different, as one can imagine. For the first path this length can easily be determined using Pythagorean theorem and is given by

\[ l = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2} = \sqrt{(\Delta y)^2 + (\Delta x)^2}. \quad (7.5) \]

That can also be put in the form

\[ l = \Delta x \sqrt{\left( \frac{\Delta y}{\Delta x} \right)^2 + 1} = \Delta y \sqrt{\left( \frac{\Delta x}{\Delta y} \right)^2 + 1}. \quad (7.6) \]
7.1. INTEGRATION IN CARTESIAN COORDINATES

For the second path, since the slope is not constant, one needs to divide the path into infinitesimal intervals, \( \Delta l_i \), find the lengths using Pythagorean theorem, and sum it up all to find all. This means

\[
\ell = \sum_i \lim_{\Delta l_i \to 0} \Delta l_i = \sum_i \lim_{\Delta x_i \to 0} \Delta x_i \sqrt{\left(\frac{\Delta y_i}{\Delta x_i}\right)^2 + 1} = \sum_i \lim_{\Delta y_i \to 0} \Delta y_i \sqrt{\left(\frac{\Delta x_i}{\Delta y_i}\right)^2 + 1}.
\]  

(7.7)

Noting that, generally, integration of a function \( f(x) \) is defined as

\[
\sum_i \lim_{\Delta f_i \to 0} \Delta f_i \to \int_1^2 df_i.
\]

(7.8)

and differentiation as

\[
\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x}
\]

(7.9)

one can write

\[
\ell = \int_1^2 dl = \int_{x_1}^{x_2} \sqrt{\left(\frac{dy}{dx}\right)^2 + 1} \, dx = \int_{y_1}^{y_2} \sqrt{1 + \left(\frac{dx}{dy}\right)^2} \, dy.
\]

(7.10)

This is the arc-length (or the total distance) for any path one can follow to go from \( P_1 \) to \( P_2 \).

B. Double integral - Area

Let’s add two points \( P_3 \) and \( P_4 \) with coordinates \((x_1, y_2)\) and \((x_2, y_1)\), respectively, to form the rectangular region shown in Fig. 7.1.1. The area of this rectangle depends on the length \( \ell \) (the distance between \( P_1 \) and \( P_4 \)) and the width \( w \) (the distance between \( P_1 \) and \( P_3 \))

\[
A = \ell w.
\]

(7.11)
The points along the length of the rectangle, \( l \), are defined by the equation

\[
x (y) = y_1 = \text{constant or } x (y) = y_2 = \text{constant}
\]  

(7.12)

and those along the width, \( w \), by

\[
y (x) = x_1 = \text{constant or } y (x) = x_2 = \text{constant}. 
\]  

(7.13)

There follows that

\[
\frac{dy}{dx} = 0, \frac{dx}{dy} = 0
\]  

(7.14)

Thus the length and the width of the rectangle can be expressed using the arc-length relation that can be expressed as

\[
l = \int_{x_1}^{x_2} \sqrt{\left( \frac{dy}{dx} \right)^2 + 1} \, dx = \int_{x_1}^{x_2} dx
\]  

(7.15)

and

\[
w = \int_{y_1}^{y_2} \sqrt{1 + \left( \frac{dx}{dy} \right)^2} \, dy = \int_{y_1}^{y_2} dy,
\]  

(7.16)

respectively. The area of the rectangle can then be expressed using the double (or surface) integral

\[
A = \int_{x_1}^{x_2} \int_{y_1}^{y_2} dy \, dx = \int_{x_1}^{x_2} dx \int_{y_1}^{y_2} dy,
\]  

(7.17)

or

\[
A = \int_{y_1}^{y_2} \int_{x_1}^{x_2} dy \, dx = \int_{y_1}^{y_2} dy \int_{x_1}^{x_2} dx.
\]  

(7.18)

Note that the order of integration does not bring any difference in the result. This result can easily be generalized for any arbitrary shape in which the boundaries can be defined by a well-defined real function. Consider the region bounded by the two curves shown in Fig. 7.1.1. Let the points on the lower boundary are defined by the function \( y_1 (x) \) and the upper boundary by the function \( y_1 (x) \). In this case the double integral for the area in Eq. ( ) should be expressed as

\[
A = \int_{x_1}^{x_2} \left[ \int_{y_1 (x)}^{y_2 (x)} dy \right] dx.
\]  

(7.19)

On the other hand, if the points in the lower boundary and the upper boundary are defined by the functions \( x_1 (y) \) and \( x_2 (y) \), respectively, instead (see Fig. ), the order of integration needs to be switched and the area should be expressed
7.1. INTEGRATION IN CARTESIAN COORDINATES

\[ A = x_2 \left[ \int_{x_1}^{x_2(y)} dy \right] dx, \]  

(7.20)

Note that even though the order of integrations are switched, the result should be the same.

C. Triple integral - Volume

\[ V = \iiint dxdydx. \]  

(7.21)

Example 1 Line and surface integral: Consider a wire bent into the arc of a circle of radius \( R \), as shown in the figure below.

(a) You climbed along this wire. Determine the height of the elevation that you climbed if you have moved a distance of half of the radius along the \( x \)-direction.

(b) Find the length of the wire

(c) Find the area bounded by the curve, the \( x \)- and \( y \)-axes

Solution:
(a) The height is the change in $y$ coordinate ($\Delta y$),

$$\Delta y = \int_{x_1}^{x_2} m(x) \, dx,$$

where $m(x)$ is the slope of the wire given by

$$m(x) = \frac{dy(x)}{dx}.$$

The function defining the curve shown in the figure is

$$y(x) = \sqrt{R^2 - x^2}$$

and slope becomes

$$m(x) = \frac{dy(x)}{dx} = -\frac{x}{\sqrt{R^2 - x^2}}.$$  \hspace{1cm} (7.23)

The distance traveled along the $x$-direction can then be written as

$$\Delta y = -\int_{R/2}^{0} \frac{xdx}{\sqrt{R^2 - x^2}} = (R^2 - x^2)^{1/2} \bigg|_{R/2}^{0} = \left(1 - \frac{\sqrt{3}}{4}\right) R.$$  \hspace{1cm} (7.24)
(b) Find the length of the wire is given by using the are-length in Eq. (7.10)

\[ l = \int_{x_1}^{x_2} \sqrt{\left( \frac{dy}{dx} \right)^2 + 1} \, dx. \]  

(7.25)

Using the result in Eq. (7.23), we may write

\[ l = \int_0^R \sqrt{\left( \frac{x}{\sqrt{R^2 - x^2}} \right)^2 + 1} \, dx = \int_0^R \frac{R \, dx}{\sqrt{R^2 - x^2}}. \]  

(7.26)

Introducing the transformation defined by

\[ x = R \sin (\theta) \Rightarrow dx = R \cos (\theta) \]  

(7.27)

the length of the wire is found to be

\[ l = \frac{\pi}{2} \frac{R^2 \cos (\theta)}{\sqrt{R^2 (1 - \sin^2 (\theta))}} \, d\theta = R \int_0^{\pi/2} \frac{R \pi}{2} \, d\theta = \frac{\pi}{2}. \]  

(7.28)

Note that we have expressed the limits of integration in terms of the variable \( \theta \). That means for \( x = 0 \)

\[ x = R \sin (\theta) \Rightarrow \theta = 0 \]

and \( x = R \)

\[ x = R \sin (\theta) \Rightarrow \theta = \sin^{-1} (1) = \pi/2 \]

Homework problem: Using the expression

\[ l = \int_{y_1}^{y_2} \sqrt{1 + \left( \frac{dx}{dy} \right)^2} \, dy, \]  

(7.29)

show that you will find the same result.

(c) The area is given by the surface integral

\[ A = \iint dxdy. \]  

(7.30)

Noting that the \( x- \) and \( y- \)axes are defined by the equations \( y = 0 \) and \( x = 0 \), respectively, for the area bounded and integration with respect to \( x \) first, we have

\[ A = \int_0^R \left( \int_0^{\sqrt{R^2 - y^2}} dx \right) dy = \int_0^R \sqrt{R^2 - y^2} \, dy. \]  

(7.31)
Introducing the transformation of variable defined by
\[ y = R \sin(\theta) \Rightarrow dy = R \cos \theta \, d\theta \]  
(7.32)
and noting that
\[ y = 0 \Rightarrow \theta = 0, y = R \Rightarrow \theta = \pi/2 \]  
(7.33)
we may rewrite the area as
\[ A = \int_0^{\pi/2} \int_0^R \sqrt{R^2 - R^2 \sin^2(\theta)} \, R \cos \theta \, d\theta \
= \frac{\pi R^2}{4} \]  
(7.34)

**Example 2 Volume integral:** A triangle in the x-y plane has vertices at the points \((x, y, z) = (0, 0, 0), (1, 0, 0), \) and \((0, 2, 0)\). Find the volume of space that lies above this triangle, but below the plane \(z = 2 + x + y\). All distance units are in centimeters.

**Solution:** The volume is given by
\[ V = \iiint dxdydz. \]  
(7.35)
The equation of the line joining the point \((1, 0, 0)\) and \((0, 2, 0)\) is given by
\[ \frac{y - 2}{x - 0} = \frac{2 - 0}{0 - 1} \Rightarrow y = -2x + 2 \]  
(7.36)
Then the volume is given by
\[ V = \int_{x=0}^{x=1} \left\{ \int_{y=0}^{y=-2x+2} \left[ \int_{z=0}^{z=2x+y} dz \right] dy \right\} dx. \]  
(7.37)
Integrating with respect to \(z\),
\[ V = \int_{x=0}^{x=1} \left\{ \int_{y=0}^{y=-2x+2} (2 + x + y) \, dy \right\} dx, \]  
(7.38)
with respect to \(y\)
\[ V = \int_{x=0}^{x=1} \left[ 2y + xy + \frac{y^2}{2} \right]_{y=0}^{y=-2x+2} \, dx = \int_{x=0}^{x=1} \left[ 2(-2x + 2) + x(-2x) \right] \, dx \]  
(7.39)
\[ = \int_{x=0}^{x=1} \left[ 6 - 6x \right] \, dx, \]  
and with respect to \(x\), results in
\[ V = \int_{x=0}^{x=1} \left[ 6 - 6x \right] \, dx = 6x - 3x^2 \bigg|_{x=0}^{x=1} = 3cm^2. \]  
(7.40)
7.1.2 Physical Applications

There are several physical quantities that require carrying out single and multiple integrals. Here are some examples in electrostatics and mechanics.

I. Electrostatics

A. Electrostatic Potential

The electrostatic properties of a given charge distribution can all be derived from the electrostatic potential, \( V(r) \). In an introductory physics course you have introduced to the electrostatic potential of a point charge. This electrostatic potential at a point in space described by the position vector, \( \vec{r} \), measured from a point charge with charge, \( q \), that is placed at the origin, is given by

\[
V(r) = \frac{q}{4\pi \varepsilon_0 |\vec{r}|},
\]

(7.41)

if the point charge is placed in a free space (with an electrical permittivity, \( \varepsilon_0 \)) in the absence of any other charge. When the charge is not a point charge and instead the charge is a line, surface, or volume charge with none uniform distribution, the electrostatic potential, \( V(r) \), is determined by carrying out single or multiple integrations that depends on the given charge distributions. For the given charge distribution let’s consider an infinitely small charge \( dq' \) at a point described by the position vector, \( \vec{r}' \), from the origin of the coordinate system. The corresponding infinitesimal electrostatic potential, \( dV(r) \), at a point in space described by the position vector, \( \vec{r} \), that is also measured from the origin, depends on the distance of this point from the position of the infinitesimal charge, \( dq' \). This distance \( R(r') \) is determined from the magnitude of the position vector directed from the position of \( dq' \) to the position of the point in space that we want to determine the corresponding electrostatic potential given by

\[
R(r') = |\vec{r} - \vec{r}'|.
\]

(7.42)

In view of Eq. (7.41), one can then express \( dV(r) \) as

\[
dV(r) = \frac{dq'}{4\pi \varepsilon_0 |\vec{r} - \vec{r}'|}.
\]

(7.43)

**Line charge density**: Suppose you are given a charge that is distributed on a finite length of wire with length, \( L \). Also the charge is not uniformly distributed on the wire and the charge distribution is described by a line a line charge density,

\[
\lambda(\vec{r}') = \frac{dq'}{dr'} \Rightarrow dq' = \lambda(\vec{r}') dr'.
\]

(7.44)

Using this equation, the total electrostatic potential due to the entire charge over the entire length can be determined from Eq. (7.43) using the
CHAPTER 7. INTEGRAL CALCULUS-SCALAR FUNCTIONS

line integral

\[ V(r) = \int dq' \left| \frac{1}{4\pi \varepsilon_0} \int \frac{\lambda(\vec{r}')}{|\vec{r} - \vec{r}'|} \, dr' \right| \]

\[ = \frac{1}{4\pi \varepsilon_0} \int L \frac{\lambda(\vec{r}')}{|\vec{r} - \vec{r}'|} \, dr' \quad (7.45) \]

Surface charge density: Any external charge that one may put on a metallic ball is distributed only on the surface of the sphere no matter the magnitude of the charge. This charge distribute itself uniformly or none uniformly depending on the absence or presence of charges in surrounding regions of the sphere. Let’s consider the general case none uniform distribution that can be described by a surface charge density,

\[ \sigma(\vec{r}') = \frac{dq'}{da'} \Rightarrow dq' = \sigma(\vec{r}') da' \quad (7.46) \]

where \( da' \) is the infinitesimal area over which the infinitesimal charge \( dq' \) is distributed on. In this case the electrostatic potential at a point \( \vec{r} \) due to the entire charge on the surface of the sphere is given

\[ V(r) = \frac{1}{4\pi \varepsilon_0} \int \int \frac{\sigma(\vec{r}')}{|\vec{r} - \vec{r}'|} \, da'. \quad (7.47) \]

This is a surface integral that requires multiple integration. Note that the infinitesimal area, \( da' \), on the surface of the sphere with radius \( r' \), in spherical coordinates can be expressed as

\[ da' = r'^2 \sin(\theta') \, d\theta' \, d\varphi' \quad (7.48) \]

Volume charge density: Similarly, for a charge described by a volume charge density,

\[ \rho(\vec{r}') = \frac{dq'}{dv'} \Rightarrow dq' = \rho(\vec{r}') dv' \quad (7.49) \]

the electrostatic potential at a point \( \vec{r} \) for the entire volume of charge can be expressed as

\[ V(r) = \frac{1}{4\pi \varepsilon_0} \int \int \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \, dv'. \quad (7.50) \]

B. The total charge: Like the total electrostatic potential depends on the charge densities and determined by single or multiple integrations, so does the total charge. The total charge, \( Q \), distributed on a line is determined by single integral

\[ Q = \int_{Line} \lambda(\vec{r}') \, dr', \quad (7.51) \]

on a surface and volume by multiple integrations given by

\[ Q = \int \int \sigma(\vec{r}') da', \quad (7.52) \]
7.1. INTEGRATION IN CARTESIAN COORDINATES

and

\[ Q = \iiint_{\text{Volume}} \rho(\mathbf{r}') dV', \]  

(7.53)

respectively.

II. Classical Mechanics

A. Moment of Inertia, \( I \): In classical mechanics, moment of inertia is also called mass moment of inertia, rotational inertia, polar moment of inertia of mass, or the angular mass. (SI units kg-m\(^2\)). It is a property of a distribution of mass in space that measures its resistance to rotational acceleration about an axis. Newton’s first law, which describes the inertia of a body in linear motion, can be extended to the inertia of a body rotating about an axis using the moment of inertia. That is, an object that is rotating at constant angular velocity will remain rotating unless acted upon by an external torque. In this way, the moment of inertia plays the same role in rotational dynamics as mass does in linear dynamics, describing the relationship between angular momentum and angular velocity, torque and angular acceleration. The symbols \( I \) and sometimes \( J \) are usually used to refer to the moment of inertia or polar moment of inertia. For a discrete mass that consist of \( N \) entities each with mass \( m_i \), the moment of inertia is given by

\[ I = \sum_{i=1}^{N} m_i r_i^2, \]  

(7.54)

for continuum mass distribution in volume, \( V \), described by the mass density, \( \rho(\mathbf{r}) \), the moment of inertia can be determined using the volume integral

\[ I = \iiint_{V} \rho(\mathbf{r}) r^2 dV = \iiint_{V} \rho(x, y, z) r^2 dxdydz \]  

(7.55)

where \( r_i (\mathbf{r}) \) is the perpendicular distance from the axis of rotation and

\[ \rho(\mathbf{r}) = \frac{dm}{dV} = \frac{dm}{dxdydz} \]  

(7.56)

is the volume mass density in Cartesian coordinates.

B. The total mass: The total mass, \( M \), in a volume with mass density, \( \rho(\mathbf{r}) \) is given by

\[ M = \iiint_{\text{Volume}} \rho(\mathbf{r}) dxdydz. \]  

(7.57)
C. Center of Mass: The center of mass (gravity) for a given mass with uniform or nonuniform distribution is a point at which the mass can be at rotational equilibrium about an axis passing through this point. The center mass, in Cartesian coordinates, \((x, y, z)\) can be determined using

\[
\bar{x} \int dm = \int x dm, \quad \bar{y} \int dm = \int y dm, \quad \text{and} \quad \bar{z} \int dm = \int z dm. \quad (7.58)
\]

For a mass distributed on a volume with volume mass density

\[
\rho (x, y, z) = \frac{dm}{dv} = \frac{dm}{dxdydz} \Rightarrow dm = \rho (x, y, z) dxdydz \quad (7.59)
\]

the center of mass can expressed as

\[
\bar{x} = \frac{\int \int \int x \rho (x, y, z) dxdydz}{\int \int \int \rho (x, y, z) dxdydz}, \quad \bar{y} = \frac{\int \int \int y \rho (x, y, z) dxdydz}{\int \int \int \rho (x, y, z) dxdydz}, \quad \bar{z} = \frac{\int \int \int z \rho (x, y, z) dxdydz}{\int \int \int \rho (x, y, z) dxdydz}. \quad (7.60)
\]

If mass is distributed along a line on the x-axis with linear mass density, we have

\[
\lambda (x) = \frac{dm}{dx} \Rightarrow dm = \lambda (x) dx \quad (7.62)
\]

and the center of mass becomes

\[
\bar{x} = \frac{\int \lambda (x) x dx}{\int dm}, \quad \bar{y} = 0, \quad \text{and} \quad \bar{z} = 0. \quad (7.63)
\]

Similarly for surface mass distribution on the x-y plane, with surface mass density

\[
\sigma (x, y) = \frac{dm}{da} = \frac{dm}{dxdy} \Rightarrow dm = \sigma (x, y) dxdy \quad (7.64)
\]

the center of mass is given by

\[
\bar{x} = \frac{\int \int x \sigma (x, y) dxdy}{\int \int \sigma (x, y) dxdy}, \quad \bar{y} = \frac{\int \int y \sigma (x, y) dxdy}{\int \int \sigma (x, y) dxdy}, \quad \text{and} \quad \bar{z} = 0. \quad (7.65)
\]

The next examples illustrates the application of single and multiple integrations to determine the physical quantities described.

**Example 3** A thin rod of length \(L\) lies along the positive x-axis with one end at the origin. It has a linear charge density given by \(\lambda (x) = Ax\), where \(A\) is a constant.
(a) What are the MKS units of the constant $A$?

(b) Find an expression for $A$ if the total charge on the rod is $Q$.

(c) Point $P$ is located a distance $b$ from the origin along the negative x-axis. Find an expression for the electrostatic potential at the point $P$, $V(r)$.

**Solution:**

(a) the unit must be the unit of charge/length$^2 = C/m^2$

(b) The total charge for a line charge distribution is given by

$$Q = \int_{\text{line}} \lambda(r') dr' = \int_0^L Ax dx = \frac{AL^2}{2} \Rightarrow A = \frac{2Q}{L^2} \quad (7.66)$$

(c) For a line charge distribution, the potential is given by

$$V(r) = \frac{1}{4\pi\epsilon_0} \int_0^L \frac{\lambda(r')}{|r-r'|} dr' \quad (7.67)$$

Using the charge density $\lambda(r') = Ax'$ and

$$\left|\vec{r} - \vec{r}'\right| = |b(-\hat{x}) - x'(\hat{x})| = b + x'$$

one can write

$$V(r) = \frac{A}{4\pi\epsilon_0} \int_0^L \frac{x'}{b + x'} dx'. \quad (7.68)$$

Noting that

$$\frac{x'}{b + x'} = 1 - \frac{b}{b + x'}, \quad (7.69)$$

one can put the integral in the form

$$V(r) = \frac{A}{4\pi\epsilon_0} \int_0^L \left[1 - \frac{b}{b + x'}\right] dx' = \frac{A}{4\pi\epsilon_0} \left[\int_0^L dx' - b \int_0^L \frac{dx'}{b + x'}\right] \quad (7.70)$$

which leads to

$$V(r) = \frac{A}{4\pi\epsilon_0} \left[L - b \ln(b + x)\right] \bigg|_{x=0}^{x=L} = \frac{A}{4\pi\epsilon_0} \left[L - b \ln \left(\frac{b + L}{b}\right)\right]. \quad (7.71)$$
Using the result in part b for \( A \), we may express the electrostatic potential as

\[
V(r) = \frac{Q}{2\pi\varepsilon_0 L} \left[ 1 - \frac{b}{L} \ln \left( \frac{b+L}{b} \right) \right]. \tag{7.72}
\]

**Example 4** A semi-circular plate of radius \( R \) is arranged in the upper-half x-y plane with its center-of-curvature at the origin. The plate has a surface-charge density given by \( \sigma(x, y) = Ax^2y \), where \( A \) is a constant.

(a) What are the MKS units of the constant \( A \)?

(b) What is the total charge on the plate in terms of \( A \)?

(c) Re-do part (b), switching the order of integration.

**Solution:**

(a) The MKS unit of \( A \) must be the unit of \( \text{charge per length} \), since we must get a unit of \( \text{charge per length}^2 \) for the surface charge density \( \sigma(x, y) \) after we multiply \( A \) by \( x^2y \).

(b) The total charge is given by surface integral

\[
Q = \int_{\text{area}} \sigma(r') da' = \int_{x_1}^{x_2} \left[ \int_{y_1(x')} \sigma(x', y') dy' \right] dx' \tag{7.73}
\]

Using equation of a circle

\[
x'^2 + y'^2 = R^2 \tag{7.74}
\]

the equation of curved part of the semi-circle in the figure can be written as

\[
y'(x') = \sqrt{R^2 - x'^2}. \tag{7.75}
\]
For the area shown in the figure the limits of integration would then be \(x_1 = -R\) to \(x_2 = R\) and \(y_1'(y') = 0\) to \(y_2'(y') = \sqrt{R^2 - y'^2}\) and the integral becomes

\[
Q = A \int_{-R}^{R} \left[ x'^2 \left( \int_{0}^{y'} y' \, dy' \right) \right] \, dx',
\]  
(7.76)

where we used the surface charge density

\[
\sigma(x', y') = Ax'^2 y'.
\]  
(7.77)

Upon carrying out the integration with respect to \(y\) first, we have

\[
Q = A \int_{-R}^{R} x'^2 \left[ \frac{y'^2}{2} \sqrt{R^2 - x'^2} \right] \, dx = A \frac{2}{3} \int_{-R}^{R} \left[ R^2 x'^2 - x'^4 \right] \, dx,
\]  
(7.78)

so that integration with respect to \(x'\) leads to

\[
Q = A \frac{2}{3} \left[ \frac{R^2 x'^3}{3} - \frac{x'^5}{5} \right]_{-R}^{R} = A \left[ \frac{R^5}{3} - \frac{R^5}{5} \right] = \frac{2}{15} A R^5
\]  
(7.79)

(c) If we switch the order of integration, we have

\[
Q = \int_{\text{area}} \sigma(r')\, da' = \int_{y_1(x_1(y'))}^{y_2(x_2(y'))} \int_{y_1(x)}^{y_2(x)} |\sigma(x', y')| \, dy' \, dx'
\]  
(7.80)

Referring to the figure, the limits of integration would then be \(y_1 = 0\) to \(y_2 = R\) and \(x_1(y') = -\sqrt{R^2 - y'^2}\) to \(x_2(y') = \sqrt{R^2 - y'^2}\) and the integral becomes

\[
Q = A \int_{0}^{R} y' \left[ \int_{-\sqrt{R^2 - y'^2}}^{\sqrt{R^2 - y'^2}} x'^2 \, dx' \right] \, dy' = A \int_{0}^{R} y' \left[ \frac{x'^3}{3} \left| \sqrt{R^2 - y'^2} \right| \right] \, dy'
\]

\[
= A \int_{0}^{R} y' \left[ \frac{2}{3} \left( R^2 - y'^2 \right)^{\frac{3}{2}} \right] \, dy'
\]  
(7.81)

The above integration can be carried out using transformation of variables. If we introduce

\[
u = (R^2 - y^2) \Rightarrow -\frac{du}{2} = ydy,
\]  
(7.82)

one can write

\[
\int y \left[ \frac{2}{3} \left( R^2 - y^2 \right)^{\frac{3}{2}} \right] \, dy = -\int \frac{u^{\frac{3}{2}}}{3} \, du = -\frac{2}{15} u^{\frac{5}{2}}
\]  
(7.83)
Therefore, the charge becomes

\[ Q = A \int_{0}^{R} \left[ \frac{2\left(R^2 - y^2\right)^{\frac{3}{2}}}{3} \right] dy = -\frac{2A}{15} \left(R^2 - y^2\right)^{\frac{3}{2}} \bigg|_{0}^{R} = \frac{2AR^5}{15} \] (7.84)

**Example 5** A right-circular cylinder has a height \( H \), a radius \( R \), and a mass \( M \) distributed uniformly throughout its volume. Find the moment of inertia of the cylinder about its axis of symmetry.

**Solution:** Since the mass is distributed uniformly in the cylinder, the volume density, \( \rho (r) \), is a constant given by

\[ \rho (r) = \frac{dm}{dv} = \frac{M}{V} = \frac{M}{\pi R^2 H}. \] (7.85)

The moment of inertia about the axis of symmetry is an axis that passes through the center of the cylinder. For a cylinder center about the z-axis, we can take the z-axis as the center of symmetry. Thus using Eq. (7.55), the moment of inertia can be expressed using a triple integral in Cartesian coordinates as

\[ I = \iiint_V \rho (r) r^2 dx dy dz = \frac{M}{\pi R^2 H} \iiint_V r^2 dx dy dz. \] (7.86)

Noting that \( r \) is perpendicular distance of the mass, \( dm \), (in the volume, \( dv = dx dy dz \), that is centered at a point with coordinates \((x, y, z)\)) from the axis of symmetry \((z\text{-axis})\), we can write

\[ r = \sqrt{x^2 + y^2}, \] (7.87)

where

\[ 0 \leq r \leq R. \] (7.88)

Furthermore, since the cylinder, with height \( H \), is centered about the z-axis,

\[ -\frac{H}{2} \leq z \leq \frac{H}{2}. \] (7.89)

Thus the moment of inertia can be expressed as

\[ I = \frac{M}{\pi R^2 H} \int_{z_1 = -H/2}^{z_2 = H/2} \left[ \int_{y_2 = R}^{y_2 = \sqrt{R^2 - y^2}} \left[ \int_{x_2 = \sqrt{R^2 - y^2}}^{x_2 = R} (x^2 + y^2) dx \right] dy \right] dz. \] (7.90)

Since the integrand is independent of \( z \) we can carry out integration with respect to \( z \) followed by integration with respect to \( y \). This leads to

\[ I = \frac{M}{\pi R^2 H} \left[ \int_{-R}^{R} \left( \frac{x^3}{3} + y^2 x \right) \left|_{-\sqrt{R^2 - y^2}}^{\sqrt{R^2 - y^2}} \right] dy \right] H \]

\[ = \frac{M}{\pi R^2} \int_{-R}^{R} 2 \left( \frac{R^2 + 2y^2}{3} \right) \sqrt{R^2 - y^2} dy. \] (7.91)
7.1. INTEGRATION IN CARTESIAN COORDINATES

Introducing the transformation of variable defined by

\[ y = R \sin(\theta) \Rightarrow dy = R \cos(\theta) \, d\theta \]  (7.92)

we may write

\[ I = \frac{M}{\pi} \int_{-\pi/2}^{\pi/2} 2 \left( \frac{R^2 + 2R^2 \sin^2(\theta)}{3} \right) \cos^2(\theta) \, d\theta \]

\[ = \frac{2MR^2}{3\pi} \int_{-\pi/2}^{\pi/2} \left[ \cos^2(\theta) + 2\sin^2(\theta) \cos^2(\theta) \right] \, d\theta. \]  (7.93)

Note that we have re-expressed the limits of integration in terms of \( \theta \) (i.e. \(-R \leq x \leq R \Rightarrow -\pi/2 \leq \theta \leq \pi/2\)). Using the result obtained by Mathematica (which you can also easily show),

\[
\text{In[1]} = I_1 = \int_{-\pi/2}^{\pi/2} \cos[\theta]^2 \, d\theta
\]

\[
\text{Out[1]} = \frac{\pi}{2}
\]

\[
\text{In[2]} = I_1 = \int_{-\pi/2}^{\pi/2} \cos[\theta]^2 \sin[\theta]^2 \, d\theta
\]

\[
\text{Out[2]} = \frac{\pi}{8}
\]

the integral becomes

\[
\int_{-\pi/2}^{\pi/2} \left[ \cos^2(\theta) + 2\sin^2(\theta) \cos^2(\theta) \right] = \frac{3\pi}{4}. \]  (7.94)

The moment of inertia is then found to be

\[ I = \frac{MR^2}{2}. \]  (7.95)

Example 6  Consider a wire of constant linear mass density \( \lambda \) bent into the arc of a circle of radius \( R \), as shown in the figure below.

(a) What is the total mass of the wire?

(b) Find the wire’s center-of-mass.

(c) Find the area under the arc of the wire.

(d) The wire is now rotated about the x-axis. Find the volume inside the surface thus generated.

(e) Exercise: Find the surface area of this curved surface of rotation.
Solution:

(a) The total mass of the wire is given by

\[ M = \int dm \]  

(7.96)

For uniformly distributed linear mass over curved length the linear mass density, \( \lambda (l) \), given by

\[ \lambda (l) = \frac{dm}{dl} = \lambda \Rightarrow dm = \lambda dl \]

the total mass becomes

\[ M = \lambda \int dl = \lambda l \]  

(7.97)

Using the result in Example 1(b) for the arc-length of the curve

\[ l = \frac{R\pi}{2}, \]

the total mass becomes

\[ M = \frac{\lambda R\pi}{2}. \]  

(7.98)

(b) The mass is distributed uniformly on a curved line on the x-y plane. Thus coordinates of the center of mass are given by

\[ \bar{x} \int dm = \int xdm, \bar{y} \int dm = \int ydm, \text{ and } \bar{z} = 0 \]  

(7.99)

Noting that from the result in (a)

\[ M = \int dm = \frac{\lambda R\pi}{2} \]  

(7.100)
The \( x \)- and \( y \)-coordinates of the center of mass can be expressed as

\[
\bar{x} = \frac{\int x \, dm}{M} = \frac{2}{\lambda R \pi} \int x \, dm = \frac{2}{R \pi} \int x \, dl
\]

and

\[
\bar{y} = \frac{\int y \, dm}{\int dm} = \frac{2}{\lambda R \pi} \int y \, dm = \frac{2}{R \pi} \int y \, dl,
\]

where we replaced \( dm = \lambda dl \) and take into account that the density is constant. From Eq. (7.10), we note that

\[
dl = \sqrt{\left(\frac{dy}{dx}\right)^2 + 1} \, dx = \sqrt{1 + \left(\frac{dx}{dy}\right)^2} \, dy
\]

we may write

\[
\bar{x} = \frac{2}{R \pi} \int_0^R x \sqrt{\left(\frac{dy}{dx}\right)^2 + 1} \, dx,
\]

and

\[
\bar{y} = \frac{2}{R \pi} \int_0^R y \sqrt{1 + \left(\frac{dx}{dy}\right)^2} \, dy.
\]

The curve is defined by the equation of a circle

\[
x^2 + y^2 = R^2 \Rightarrow y = \sqrt{R^2 - y^2} \text{ or } x = \sqrt{R^2 - y^2},
\]

we chose the positive roots for both \( x \) and \( y \) since the in the first quadrant both are positives. Using the expressions for \( x \) and \( y \) we have

\[
\frac{dx}{dy} = -\frac{y}{\sqrt{R^2 - y^2}} \quad \frac{dy}{dx} = -\frac{x}{\sqrt{R^2 - x^2}}
\]

so the one can write

\[
\bar{x} = \frac{2}{R \pi} \int_0^R x \sqrt{\left(\frac{x}{\sqrt{R^2 - x^2}}\right)^2 + 1} \, dx = \frac{2}{\pi} \int_0^R \frac{xdx}{\sqrt{R^2 - x^2}} x \, dx,
\]

and

\[
\bar{y} = \frac{2}{R \pi} \int_0^R y \sqrt{\left(\frac{y}{\sqrt{R^2 - y^2}}\right)^2 + 1} \, dx = \frac{2}{\pi} \int_0^R \frac{ydy}{\sqrt{R^2 - y^2}} y \, dx.
\]

Upon carrying out the integrations,

\[
\bar{x} = \frac{2}{\pi} \left(-\sqrt{R^2 - x^2}\right)^R_0, \quad \bar{y} = \frac{2}{\pi} \left(-\sqrt{R^2 - x^2}\right)^R_0.
\]
and one finds

\[ \bar{x} = \bar{y} = \frac{2R}{\pi} \]  

(7.110)

The center of mass is \((\frac{2R}{\pi}, \frac{2R}{\pi}, 0)\).

(c) The area can be determined using Eqs. (7.19)

\[ A = \int_{x_1}^{x_2} \left[ \int_{y_1(x)}^{y_2(x)} dy \right] dx. \]  

(7.111)

or (7.20)

\[ A = \int_{x_1}^{x_2} \left[ \int_{y_1(y)}^{y_2(y)} dy \right] dx. \]  

(7.112)

Here we chose to use Eq. (7.19). For the boundaries shown in the figure, we have

\[ x_1(y) = 0, x_2(y) = \sqrt{R^2 - x^2} \]  

(7.113)

and

\[ x_1 = 0, x_2 = R. \]  

(7.114)

so that the area becomes

\[ A = \int_{0}^{R} \left[ \int_{0}^{\sqrt{R^2 - x^2}} dy \right] dx = \int_{0}^{R} \sqrt{R^2 - x^2} dx. \]  

(7.115)

Introducing the transformation of variable defined by

\[ x = R \sin (\theta) \Rightarrow dx = R \cos \theta d\theta, \]  

(7.116)

we may write

\[ A = \int_{0}^{\pi/2} \sqrt{R^2 - R^2 \sin^2 \theta} R \cos \theta d\theta = R^2 \int_{0}^{\pi/2} \cos^2 \theta d\theta. \]  

(7.117)

Employing the relation

\[ \cos (2\theta) = \cos^2 (\theta) - \sin^2 (\theta) \Rightarrow \cos^2 (\theta) = \frac{\cos (2\theta) + 1}{2} \]  

(7.118)

the area is found to be

\[ A = R^2 \int_{0}^{\pi/2} \left[ \frac{\cos 2\theta + 1}{2} \right] d\theta = \frac{\sin 2\theta}{4} + \frac{\theta}{2} \Big|_{0}^{\pi/2} = \frac{\pi R^2}{4}. \]  

(7.119)

Note that the limits of integration in terms of \(\theta = \sin^{-1} (x/R)\),

\[ x = 0 \Rightarrow \theta = 0, \text{ and } x = R \Rightarrow \theta = \frac{\pi}{2} \]  

(7.120)

are used in the integration.
(d) The rotation of the wire about the x-axis creates a hemisphere of radius, $R$, shown in the figure. A point on the surface of the sphere with coordinates $(x, y, z)$ is defined by the equation of a sphere

$$x^2 + y^2 + z^2 = R^2 \Rightarrow x = \sqrt{R^2 - (y^2 + z^2)} \tag{7.121}$$

and you can determine the volume by evaluating the integral

$$V = \int_{-R}^{R} \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \int_{0}^{\sqrt{R^2-(y^2+z^2)}} dz \, dy \, dx. \tag{7.122}$$

But here we will use a short cut. If we consider the infinitesimal volume to be a cylinder of radius $y$ and thickness (i.e. height) $dx$, we can write

$$dv = \pi y^2 dx \tag{7.123}$$

where $y$ can be related to the radius of the sphere by

$$y = \sqrt{R^2 - x^2}. \tag{7.124}$$

The volume will then be

$$V = \int_{0}^{R} \pi y^2 dx = \pi \int_{0}^{R} (R^2 - x^2) dx = \pi \left( R^3 - \frac{R^3}{3} \right) = \frac{2\pi R^3}{3}. \tag{7.125}$$

### 7.2 Integration in curvilinear coordinates

We now consider the methods of integrations in curvilinear coordinates. We first consider one and two dimensional curvilinear coordinates. For three-dimensional
coordinates, integrations in cylindrical and spherical coordinates will be considered. For problems with curved geometries, the simplicity of carrying out the integrations in curvilinear instead of Cartesian coordinates will also be demonstrated in this section. The demonstrations will re-consider some of the problems we saw as examples in the previous section.

7.2.1 One- and two-dimensional curvilinear Coordinates

In polar coordinates a point in the x-y plane with coordinates \((x, y)\) can be described by \((r, \theta)\). These two system of coordinates are related by

\[
x = r \cos \theta, \quad y = r \sin \theta.
\]

The infinitessimal arc-length in Cartesian coordinates that we saw in the previous section

\[
dl = \sqrt{dx^2 + dy^2}
\]

in polar coordinates can then be written as

\[
dl = \sqrt{\left[d(r \cos \theta)\right]^2 + \left[d(r \sin \theta)\right]^2} = \sqrt{[\cos \theta dr - r \sin \theta d\theta]^2 + [\sin \theta dr + r \cos \theta d\theta]^2}.
\]

This can be simplified into

\[
dl = \sqrt{(\cos^2 \theta + \sin^2 \theta) dr^2 + r^2 (\sin^2 \theta + \cos^2 \theta) d\theta^2} = \sqrt{dr^2 + r^2 d\theta^2}.
\]

By factoring \(dr\) or \(r d\theta\), one can establish the relations

\[
ds = \sqrt{1 + \left(r \frac{d\theta}{dr}\right)^2} dr
\]

or

\[
ds = \sqrt{\left(\frac{1}{r} \frac{dr}{d\theta}\right)^2 + 1} r d\theta.
\]

respectively. Similarly for the infinitessimal area in Cartesian coordinates,

\[
da = dxdy,
\]

Using

\[
dx = \frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \theta} d\theta = \cos \theta dr - r \sin \theta d\theta
\]

and

\[
dy = \frac{\partial y}{\partial r} dr + \frac{\partial y}{\partial \theta} d\theta = \sin \theta dr + r \cos \theta d\theta
\]

one can write

\[
da = dxdy = (\cos \theta dr - r \sin \theta d\theta) (\sin \theta dr + r \cos \theta d\theta).
\]
There follows that
\[ da = \cos^2(\theta) r \, dr \, d\theta - \sin^2(\theta) r \, dr \, d\theta + \sin(\theta) \cos(\theta) (dr)^2 - \sin(\theta) \cos(\theta) (r \, d\theta)^2. \] (7.133)
\[ da = r \, dr \, d\theta - 2 \sin^2(\theta) r \, dr \, d\theta + \sin(\theta) \cos(\theta) \left[(dr)^2 - (r \, d\theta)^2\right]. \] (7.134)
which is given by
\[ da = r \, dr \, d\theta. \] (7.135)

**Example 7** Using curvilinear coordinates, repeat Example 6a-d

(a) The total mass of the wire is given by
\[ M = \int dm = \int \rho_L \, ds \] (7.136)
In polar coordinates this can be expressed as
\[ M = \int_0^{\pi/2} \rho_L \sqrt{\left(\frac{1}{r} \frac{dr}{d\theta}\right)^2 + 1} r \, dr \, d\theta \] (7.137)
Referring to the figure, we have
\[ r = R \Rightarrow \frac{dr}{d\theta} = 0 \] (7.138)
so that
\[ M = \int_0^{\pi/2} \rho_L R \, d\theta = \frac{\rho_L R \pi}{2} \] (7.139)

(b) The area is given by
\[ A = \int da = \int \int dx \, dy \] (7.140)
Noting that the infinitesimal area \( da \) can be written as
\[ da = r \, dr \, d\theta \] (7.141)
we may express the area as
\[ A = \int_0^{\pi/2} \int_0^R r \, dr \, d\theta = \frac{\pi R^2}{4} \] (7.142)

**Example 8** (This is Example 4b) A semi-circular plate of radius \( R \) is arranged in the upper-half x-y plane with its center-of-curvature at the origin. The plate has a surface-charge density given by \( \sigma(x, y) = A x^2 \, y \), where \( A \) is a constant.

Determine the total charge on the plate using polar coordinates.

**Solution:**
7.2.2 3-D Curvilinear coordinates- cylindrical

A point \( p \) in space described by the Cartesian coordinates \((x, y, z)\) can be described in cylindrical coordinates by \((r, \varphi, z)\). These coordinates are related by

\[
x = r \cos(\varphi), \quad y = r \sin(\varphi), \quad z = z
\]  

(7.143)

In cylindrical coordinates an infinitesimal volume \( dv \) can be expressed as

\[
dv = dr \ (r d\varphi) \ dz = r dr d\varphi dz
\]  

(7.144)

infinitesimal area \( da \) on the curved surface

\[
da = dz \ (r d\varphi) = rdz d\varphi
\]  

(7.145)

on the planar surface

\[
da = (dr) \ (r d\varphi) = r dr d\varphi
\]  

(7.146)
7.2. INTEGRATION IN CURVILINEAR COORDINATES

Moment of Inertia: The moment of inertia of an infinitesimal mass $dm$ in the infinitesimal volume $dv$ about a given axis of rotation is given by

$$I = \int r_\perp dm = \int r_\perp \rho dv$$  \hspace{1cm} (7.147)

where $r_\perp$ is the perpendicular distance of the mass $dm$ from the given axis of rotation and $\rho$ is the mass density. In cylindrical coordinates this can be expressed as

$$I = \int r_\perp dm = \int \int r_\perp \rho r dr d\phi dz$$  \hspace{1cm} (7.148)

Example 9 (This is the same problem considered in Example 17.4) A right-circular cylinder has a height $H$, a radius $R$, and a mass $M$ distributed uniformly throughout its volume. Find the moment of inertia of the cylinder about its axis of symmetry using cylindrical coordinates.

Solution: Moment of inertia of a solid object is given by

$$I = \int \int \int r^2 dm = \int \int \int r^2 \rho (x, y, z) dxdydz$$  \hspace{1cm} (7.149)

The mass is distributed uniformly and therefore the density is constant

$$I = \rho \int \int \int r^2 dxdydz = \frac{M}{V} \int \int \int r^2 dxdydz = \frac{M}{\pi R^2 H} \int \int \int r^2 dxdydz$$  \hspace{1cm} (7.150)

The axis of symmetry is the $z$-axis if we assume the cylinder is vertical. Then we may express the inertia about this axis as

$$I = \int \int \int r^2 dm = \frac{M}{\pi R^2 H} \int \int \int r^2 r dr d\phi dz$$  \hspace{1cm} (7.151)

$$I = \frac{M}{\pi R^2 H} \int_0^H \int_0^R r^3 dr d\phi dz = \frac{1}{2} MR^2$$  \hspace{1cm} (7.152)

Example 10 Using Cylindrical coordinates prove that the volume of a right-circular cone of base radius $R$ and height $H$ to be

$$V = \frac{\pi R^2 H}{3}.$$  \hspace{1cm} (7.153)

Solution: In cylindrical coordinates an infinitesimal volume $dv$ is given by

$$dv = r dr d\phi dz \Rightarrow V = \int \int \int r dr d\phi dz$$  \hspace{1cm} (7.154)
Noting that for a right circular cone, we have
\[ 0 < \varphi < 2\pi, 0 < r < \frac{(H - z)R}{H}, 0 < z < H, \] (7.155)
and the volume can be written as
\[ V = \int_0^{2\pi} \left[ \int_0^H \left( \int_0^{\frac{(H - z)R}{H}} r \, dr \right) \, dz \right] \, d\varphi. \] (7.156)

Upon integrating with respect to \( \varphi \) and then \( r \), one finds
\[ V = 2\pi \int_0^H \frac{r^2}{2} \left[ \frac{(H - z)R}{H} \right] \, dz = \frac{\pi R^2}{H^2} \int_0^H (H^2 - 2Hz + z^2) \, dz. \] (7.157)

There follows that
\[ V = \frac{\pi R^2}{H^2} \left[ \left. H^2z - Hz^2 + \frac{z^3}{3} \right|_0^H \right] = \frac{\pi R^2H}{3}. \] (7.158)

### 7.2.3 3-D Curvilinear Coordinate Systems: spherical Coordinates

A point \( p \) in space described by the Cartesian coordinates \((x, y, z)\) can be described in spherical coordinates by \((r, \theta, \varphi)\). These coordinates are related by
\[ x = r \cos (\varphi) \sin (\theta), \quad y = r \sin (\varphi) \sin (\theta), \quad z = r \cos (\theta) \]

In spherical coordinates an infinitesimal volume \( dv \) can be expressed as
\[ dv = dr (r d\theta) (r \sin (\theta) \, d\varphi) = r^2 dr \sin (\theta) \, d\theta d\varphi \] (7.159)

infinitesimal area \( da \) on the surface
\[ da = (r d\theta) (r \sin (\theta) \, d\varphi) = r^2 \sin (\theta) \, d\theta d\varphi \] (7.160)
Example 11 A thin spherical dielectric shell of radius $R$ has a surface-charge density given by
\[ \sigma(\theta, \varphi) = \sigma_0 \sin(\theta) \cos\left(\frac{\varphi}{4}\right) \] (7.161)
where $\sigma_0$ is a constant. Find the total charge.

Solution: An infinitesimal area on the surface of a sphere of radius $R$ is given by
\[ da = R^2 \sin(\theta) \, d\theta d\varphi \] (7.162)
Thus the total surface charge given by
\[ Q = \int \int \sigma(\theta, \varphi) \, da \] (7.163)
can be expressed as
\[ Q = \int \int \sigma_0 \sin(\theta) \cos\left(\frac{\varphi}{4}\right) R^2 \sin(\theta) \, d\theta d\varphi \] (7.164)
For a full sphere the limits of integration for $\theta$ is $(0 \leq \theta \leq \pi)$ and for $\varphi$ is $(0 \leq \varphi \leq 2\pi)$. Thus the total charge can be determined by evaluating the integral
\[
Q = \sigma_0 R^2 \int_0^{2\pi} \int_0^\pi \sin^2(\theta) \cos\left(\frac{\varphi}{4}\right) \, d\theta d\varphi \\
= \sigma_0 R^2 \left( \int_0^\pi \sin^2(\theta) \, d\theta \right) \left( \int_0^{2\pi} \cos\left(\frac{\varphi}{4}\right) \, d\varphi \right) \] (7.165)
Recalling that
\[
\int_0^\pi \sin^2(\theta) \, d\theta = \theta - \frac{\sin 2\theta}{4} \bigg|_0^\pi = \frac{\pi}{2}
\]
and
\[
\int_0^{2\pi} \cos \left(\frac{\varphi}{4}\right) \, d\varphi = 4\sin \frac{\varphi}{4} \bigg|_0^{2\pi} = 4
\]
The total charge is found to be
\[
Q = 2\pi \sigma_0 R^2.
\]

Example 12 A zone of a sphere is defined to be that section of the sphere that lies between two parallel planes intersecting the sphere. Consider two planes separated by a distance \(h\) that intersect a sphere of radius \(R\). Prove that the surface area of a zone cut out by these two planes is \(A = 2\pi Rh\). Prove this somewhat surprising result (surprising in that it does not depend on where the slice of the sphere is taken—toward the equator of the sphere or closer to the top or bottom!)

Solution: An infinitesimal area \(da\) on the surface of the sphere with radius \(R\) is given by
\[
da = R^2 \sin(\theta) \, d\theta d\varphi.
\]
Then the surface area of the zone is given by
\[
A = \int da = \int_0^{2\pi} \int_{\theta_1}^{\theta_2} R^2 \sin(\theta) \, d\theta d\varphi = -2\pi R^2 \cos(\theta) \bigg|_{\theta_1}^{\theta_2}
= 2\pi R^2 \left[\cos(\theta_1) - \cos(\theta_2)\right].
\]
where the limits of integration, \( \theta_1 \) and \( \theta_2 \) represent the angular positions of the two planes forming the zone. Suppose in Cartesian coordinates, the upper plane is defined by \( z = z_1 \) and the bottom by \( z = z_1 \), we can easily see that

\[
\cos(\theta_1) = \frac{z_1}{R}, \quad \cos(\theta_2) = \frac{z_2}{R}.
\] (7.171)

There follow that

\[
R[\cos(\theta_1) - \cos(\theta_2)] = z_1 - z_2 = h,
\] (7.172)

where \( h \) is the height of the zone. Thus the area of the surface area of the zone of a sphere becomes

\[
A = 2\pi Rh,
\] (7.173)

which is indeed independent of the position of the two planes forming the zone.

### 7.3 Homework Assignment 16

1. Consider an area on the x-y plane bounded by the y-axis, the line defined by the equation, \( x = 2 \), and a curve defined by the equation, \( y = \sqrt{x} \). Using integration in Cartesian coordinates, find
   
   (a) the length of the curved side of the area,
   
   (b) the area.

2. The area in the previous problem is rotated about the y-axis to form a volume with curved surface. Using integration in Cartesian coordinates, find the surface area for
   
   (a) the curved surface,
   
   (b) the plane surface.

3. For an area \( (A) \) bounded by the x-axis \( (y = 0) \), \( y = \sin(x) \), and \( x = \pi/2 \) evaluate the integral

\[
I = \int \int_A 6y^2 \cos x \, dx \, dy.
\]

4. For the area \( (A) \) bounded by the parabola \( y = x^2 \) and the straight line \( 2x - y + 8 = 0 \), evaluate the integral

\[
I = \int \int_A x \, dx \, dy = 36
\]

5. Using double integrals determine the volume bounded above the square plane formed by the vertices \((0, 0), (2, 0), (0, 2), \) and \((2, 2)\) and below the plane \( z = 8 - x + y \).
7.4 Homework Assignment 17

1. A rope of length, \( L = 10 \text{ cm} \), is stretched along the positive y-axis with one end at the origin. The linear mass density, \( \lambda(y) \), over this rope increases uniformly from 4 \( \text{gm/cm} \) at its one end at the origin to 24 \( \text{gm/cm} \) at the other end. Find

(a) the linear mass density,
(b) the total mass of the rope,
(c) the center of mass,
(d) the moment of inertia about an axis passing through the center of mass,
(e) the moment of inertia about an axis passing through the heavy end.

2. A solid disk with radius, \( r = a \), and negligible thickness is sitting on the x-y plane centered at the origin. It has a uniform surface mass density, \( \sigma_0 \). Using integration in Cartesian coordinates, find

(a) the circumference of the circle,
(b) the area of the disk,
(c) the centroid (or center of mass) for the portion of the disk in the first quadrant,
(d) the moment of inertia of the disk about an axis along the diameter of the disk,
(e) the centroid for the portion of the side of the disk in the first quadrant,

3. The surface charge density (\( \sigma \)) on the area bounded by the x-axis \( (y = 0) \) and curve \( (y = 16 - x^2) \) is proportional to \( y \). That means

\[ \sigma(y) = ky. \]

Find the total charge (\( Q \)) on the area.

4. A light with an intensity, \( I_0 = 42 \text{ J/m}^2 \), is incident on a square silvered mirror sitting on the x-y plane. The square has side length 2 \( m \) and is centered about the origin. The fraction of the incident light that is reflected at a point on the mirror \( (x, y) \) is \( (x - y)^2 / 4 \). In other words, the intensity of the reflected light at the point is

\[ I_R(x, y) = I_0 (x - y)^2 / 4 \]

Find the fraction of the reflected light from the entire mirror.

5. Using double integrals determine the volume formed below the surface defined by \( z = y(x + 2) \) and above the surface bounded by \( x + y = 0, y = 1, \) and \( y = \sqrt{x} \).
7.5 Homework Assignment 18

1. Repeat Problem 2 in Homework Assignment 16 using Polar coordinates.

1. Consider the volume of the space bounded by the curved surface of the cylinder defined $x^2 + y^2 = 4$, $z = 2x^2 + y^2$, and the x-y plane. Find the volume of the cylinder using integrations in cylindrical coordinates.

3. Consider a solid sphere of radius, $r = a$ and mass, $M$. The mass is distributed uniformly over the volume of the sphere. Using integrations in spherical coordinates, find

(a) The centroid of a solid half ball,

(b) the moment of inertia of the of the sphere about an axis passing through the diameter (along the z-axis)

(c) the volume of the sphere

4. In problem 3 we considered a solid sphere of radius, $r = a$, and mass, $M$ that is distributed uniformly over the volume of the sphere. Now let’s imagine the sphere is hollow and the same mass is distributed over the surface of the sphere. Using integrations in spherical coordinates, find

(a) the surface area of the sphere,

(b) the centroid of the curved area of the upper hemisphere,

(c) the moment of inertia of the of the sphere about an axis passing through the diameter (along the z-axis)

5. Consider a sphere with constant mass density, $\rho_0$ and radius $r = 2a$. Determine the gravitational attraction force on a unit mass ($m = 1$) at the center of the sphere due to the portion of the mass of the sphere occupying the volume above the plane defined by, $z = a$. 

Chapter 8

Vector Calculus

We were introduced to the basic of vector algebra in chapter 3. Here we will introduced to calculus of vector functions (differentiation and integration). But first we begin by reviewing vector products and their applications in physics.

8.1 Vector Products

Two vectors \( \mathbf{A} = A_x \hat{x} + A_y \hat{y} + A_z \hat{z} \) and \( \mathbf{B} = B_x \hat{x} + B_y \hat{y} + B_z \hat{z} \) can be multiplied in two different ways one resulting a scalar and the other resulting a third different vector.

*Dot product (Scalar product):* The dot product of \( \mathbf{A} \) and \( \mathbf{B} \) is given by

\[
\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z \tag{8.1}
\]

If the angle between the two vectors is \( \theta \) the dot product can be determined using

\[
\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}| |\mathbf{B}| \cos \theta \tag{8.2}
\]

*Cross product (vector product):* the cross product of \( \mathbf{A} \) and \( \mathbf{B} \) results in a third vector \( \mathbf{C} \) given by

\[
\mathbf{C} = \mathbf{A} \times \mathbf{B} = \det \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \tag{8.3}
\]

\[
\mathbf{C} = (A_y B_z - A_z B_y) \hat{x} + (A_z B_x - A_x B_z) \hat{y} + (A_x B_y - A_y B_x) \hat{z}. \tag{8.4}
\]

If the angle between the two vectors is \( \theta \), the magnitude of vector \( \mathbf{C} \) is given by

\[
|\mathbf{C}| = |\mathbf{A} \times \mathbf{B}| = |\mathbf{A}| |\mathbf{B}| \sin (\theta) \tag{8.5}
\]

and the direction is determined using the right-hand rule. Vector \( \mathbf{C} \) is always perpendicular to the plane formed by the two vectors.
**Triple scalar and vector products:** Consider the three vectors \( \vec{A}, \vec{B}, \) and \( \vec{C} \) in Cartesian coordinates

\[
\vec{A} = A_x \hat{x} + A_y \hat{y} + A_z \hat{z}, \quad \vec{B} = B_x \hat{x} + B_y \hat{y} + B_z \hat{z}, \quad \vec{C} = C_x \hat{x} + C_y \hat{y} + C_z \hat{z}. \tag{8.6}
\]

These three vectors can be multiplied to give a scalar or a vector. The product of these vectors that leads to a scalar is called a **triple scalar product** and is given by

\[
\vec{C} \cdot (\vec{A} \times \vec{B}) = \det \begin{vmatrix} C_x & C_y & C_z \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}.	ag{8.7}
\]

Recalling that interchanging two rows changes only make the value of the determinant negative, one can write

\[
\begin{vmatrix} C_x & C_y & C_z \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} = -\begin{vmatrix} B_x & B_y & B_z \\ A_x & A_y & A_z \\ C_x & C_y & C_z \end{vmatrix} = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}.	ag{8.8}
\]

Referring to the expression for the triple scalar product, one can easily see that

\[
\vec{C} \cdot (\vec{A} \times \vec{B}) = -\vec{B} \cdot (\vec{A} \times \vec{C}) = \vec{A} \cdot (\vec{B} \times \vec{C}). \tag{8.9}
\]

Taking into account the relation

\[
\vec{A} \times \vec{C} = -\vec{C} \times \vec{A},
\]

one can establish the relation for triple scalar product of three vectors

\[
\vec{C} \cdot (\vec{A} \times \vec{B}) = \vec{B} \cdot (\vec{C} \times \vec{A}) = \vec{A} \cdot (\vec{B} \times \vec{C}). \tag{8.10}
\]

These same three vectors can also be multiplied to yield another vector instead of a scalar. Such multiplication is known as **triple vector product** that can be expressed as

\[
\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B} \left( \vec{A} \cdot \vec{C} \right) - \vec{C} \left( \vec{A} \cdot \vec{B} \right).
\]

The expression for triple vector product can be memorized using the phrase "**BAC-CAB RULE**".

### 8.1.1 Physical applications

**Work:** Suppose a constant force, \( \vec{F} \), is applied on an object to move it for from a position \( \vec{r}_1 \) to a position \( \vec{r}_2 \) for a displacement, \( \vec{d} = \Delta \vec{r} = \vec{r}_2 - \vec{r}_1 \). The angle between the force and the displacement vectors is \( \theta \). We have seen that the work done by this force, \( W \), is given by the scalar product of the vectors \( \vec{F} \) and \( \vec{d} \),

\[
W = |\vec{F}| |\vec{d}| \cos \theta = \vec{F} \cdot \vec{d}.	ag{8.11}
\]
8.1. VECTOR PRODUCTS

For none constant force that depends on the position of the object, \( \vec{F} = \vec{F} (\vec{r}) \), the work done is determined from an integral form of the scalar product for the force and displacement vectors given by

\[
W = \int_{\vec{r}_1}^{\vec{r}_2} \vec{F} \cdot d\vec{r}.
\]  

(8.12)

**Angular and linear velocities:** Consider an object under a rotational motion. Suppose the object is rotating with a constant angular velocity, \( \vec{\omega} \), about some axis of rotation. Over a time interval, \( \Delta t = t_2 - t_1 \), the position of the object relative to the axis of rotation has changed from \( \vec{r}_1 \) to \( \vec{r}_2 \) and the linear displacement becomes, \( \vec{d} = \Delta \vec{r} = \vec{r}_2 - \vec{r}_1 \). The linear velocity, \( \vec{v} \), of the object, which is a constant for a constant angular velocity, \( \vec{\omega} \), can be determined using the vector product of the two vectors expressed as

\[
\vec{v} = \vec{\omega} \times \vec{d}.
\]  

(8.13)

However, when the angular velocity is not a constant and depends on the relative position of the object, \( \vec{\omega} = \vec{\omega} (\vec{r}) \), the average linear velocity is determined using the line integral,

\[
\vec{v} = \int_{\vec{r}_1}^{\vec{r}_2} \vec{\omega} \times d\vec{r}.
\]  

(8.14)

**Angular momentum:** Suppose the object for which we described the constant linear velocity, \( \vec{v} \), and constant angular velocity, \( \vec{\omega} \) has a mass, \( m \). This mass has a linear momentum, \( \vec{p} = m \vec{v} \). The angular momentum, \( \vec{J} \), of this object, is given by the vector product,

\[
\vec{J} = \vec{r} \times \vec{p}.
\]  

(8.15)

When the linear momentum depends on the position of the object the average angular momentum becomes

\[
\vec{J} = \int d\vec{r} \times \vec{p}.
\]  

(8.16)

**Torque:** The mechanical torque, \( \vec{\tau} \), is vector physical quantity that we often use to describe the rotational motion of an object acted by a force, \( \vec{F} \). Suppose the force acted on the object has rotated the object about an axis such that the position relative to the axis has changed from \( \vec{r}_1 \) to \( \vec{r}_2 \) for a total displacement, \( \vec{d} = \Delta \vec{r} = \vec{r}_2 - \vec{r}_1 \). For a constant force the torque about the axis of rotation is given by the vector product,

\[
\vec{\tau} = \vec{d} \times \vec{F}.
\]  

(8.17)

However, when the force is not constant and depends on the position of the object, the torque given by the integral,

\[
\vec{\tau} = \int_{\vec{r}_1}^{\vec{r}_2} d\vec{r} \times \vec{F}.
\]  

(8.18)
Replace the linear velocity, \( \vec{v} = \vec{\omega} \times \vec{r} \), in the expression for the linear momentum \( \vec{p} = m\vec{v} \), the angular momentum can be expressed as a triple vector product.

\[
\vec{J} = m \int d\vec{r} \times (\vec{\omega} \times \vec{r}) \quad (8.19)
\]

**Magnetic field and force:** Suppose an object with charge, \( q \), is moving with a constant velocity, \( \vec{v} \), in a region with a uniform magnetic field, \( \vec{B} \). The magnetic force experienced by the object is given by

\[
\vec{F}_B = q\vec{v} \times \vec{B} \quad (8.20)
\]

For a non-uniform magnetic field, \( \vec{B}(\vec{r}) \), such as the magnetic field due to a wire of length \( l \) that carries a current, \( I \), where the magnetic field is given by

\[
\vec{B}(\vec{r}) = \frac{\mu_0 I}{4\pi} \frac{\vec{l} \times \vec{r}}{r^3} \quad (8.21)
\]

the magnetic force can be expressed using a triple vector product

\[
\vec{F}_B(\vec{r}) = \frac{q\mu_0 I}{4\pi} \frac{\vec{v} \times (\vec{l} \times \vec{r})}{r^3} \quad (8.22)
\]

**Solid state Physics:** Reciprocal lattice vectors \( \{\vec{b}_i\} \)

\[
\vec{b}_i = \frac{\vec{a}_j \times \vec{a}_k}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} , \text{ where } (i, j, k) = (1, 2, 3) \quad (8.23)
\]

\( \{\vec{a}_i\} \) is called a lattice vector. It is used to describe crystal structure in solids. The reciprocal lattice vectors are useful to study waves in solids (lattice waves) known as phonons. These waves are described by the wave vector, \( \vec{k} \), expressed in terms of the reciprocal vectors.

### 8.1.2 Vectors derivatives

Next we shall see how we carry out the derivatives of a vector function. To this end, let’s consider time dependent scalar function, \( c(t) \), and two vector functions that can be expressed in Cartesian coordinates as,

\[
\vec{A}(t) = A_x(t) \hat{x} + A_y(t) \hat{y} + A_z(t) \hat{z}, \quad \vec{B}(t) = B_x(t) \hat{x} + B_y(t) \hat{y} + B_z(t) \hat{z}. \quad (8.24)
\]

Using these vectors we can establish the following derivatives with respect to time:

**Derivative of a vector:**

\[
\frac{d\vec{A}}{dt} = \frac{dA_x}{dt} \hat{x} + \frac{dA_y}{dt} \hat{y} + \frac{dA_z}{dt} \hat{z} \quad (8.25)
\]
8.1. VECTOR PRODUCTS

Derivative of a vector multiplied by a scalar:

\[ \frac{d}{dt}(c \vec{A}) = \vec{A} \frac{dc}{dt} + \frac{d\vec{A}}{dt} c \]  

(8.26)

Derivative of scalar vector product:

\[ \frac{d}{dt} (\vec{A} \cdot \vec{B}) = \vec{A} \cdot \frac{d\vec{B}}{dt} + \frac{d\vec{A}}{dt} \cdot \vec{B} \]  

(8.27)

Derivative of vector product:

\[ \frac{d}{dt} (\vec{A} \times \vec{B}) = \vec{A} \times \frac{d\vec{B}}{dt} + \frac{d\vec{A}}{dt} \times \vec{B} \]  

(8.28)

Example 8.1 The position of an object at time, \( t \), is described by the two-dimensional vector, \( \vec{r}(t) \). This vector, using Polar coordinates, \( (r, \theta) \), can be expressed as

\[ \vec{r}(t) = r \cos (\theta) \hat{x} + r \sin (\theta) \hat{y}. \]  

(8.29)

Find the general expressions for the velocity, \( \vec{v}(t) \), and acceleration, \( \vec{a}(t) \), and express your answer in Polar Coordinates.

Solution:

The velocity: \( \vec{v}(t) \), is given by

\[ \vec{v}(t) = \frac{d\vec{r}}{dt} = \frac{dr}{dt} (r \cos (\theta)) \hat{x} + \frac{dr}{dt} (r \sin (\theta)) \hat{y} \]  

(8.30)

Assuming the general case where both \( r \) and \( \theta \) are time dependent, upon differentiating, we find

\[ \vec{v}(t) = \frac{dr}{dt} \cos (\theta) \hat{x} + \frac{dr}{dt} \sin (\theta) \hat{y} + \frac{d\theta}{dt} \hat{r} \]  

(8.31)

that can put in the form

\[ \vec{v}(t) = \frac{dr}{dt} \cos (\theta) \hat{x} + \frac{dr}{dt} \sin (\theta) \hat{y} + \frac{d\theta}{dt} \hat{r}. \]  

(8.32)

Noting that in two-dimensional Polar coordinates, the unit vectors along the radial direction (i.e. along \( \vec{r} \)) is

\[ \hat{r} = \frac{d\vec{r}}{dr} \left/ \left| \frac{d\vec{r}}{dr} \right| \right. = \cos (\theta) \hat{x} + \sin (\theta) \hat{y}, \]  

(8.33)

and along the angular direction is

\[ \hat{\theta} = \frac{d\vec{r}}{d\theta} \left/ \left| \frac{d\vec{r}}{d\theta} \right| \right. = - \sin (\theta) \hat{x} + \cos (\theta) \hat{y}; \]  

(8.34)
CHAPTER 8. VECTOR CALCULUS

The instantaneous velocity can be put in takes the form

$$\mathbf{v}(t) = \frac{dr}{dt} \mathbf{\hat{r}} + \frac{d\theta}{dt} \mathbf{\hat{\theta}}. \quad (8.35)$$

In terms of the angular speed, \( \omega \),

$$\omega = \frac{d\theta}{dt} \quad (8.36)$$

and the radial speed

$$v_r = \frac{dr}{dt}, \quad (8.37)$$

the instantaneous velocity can be expressed as

$$\mathbf{v}(t) = v_r \mathbf{\hat{r}} + \omega r \mathbf{\hat{\theta}}. \quad (8.38)$$

Further more, using the equation relating the linear tangential and the angular speeds,

$$v_t = \omega r, \quad (8.39)$$

the instantaneous can also be written as

$$\mathbf{v}(t) = v_r \mathbf{\hat{r}} + v_t \mathbf{\hat{\theta}}. \quad (8.40)$$

When the magnitude of the radial displacement is constant (i.e. the motion is in a circular orbit), we have

$$v_r = \frac{dr}{dt} = 0 \quad (8.41)$$

and the instantaneous velocity becomes

$$\mathbf{v}(t) = v_t \mathbf{\hat{\theta}} = \omega r \mathbf{\hat{\theta}}. \quad (8.42)$$

This is in agreement with what we learned in into physics I. Acceleration: \( \mathbf{a}(t) \) is given by the time derivative of the instantaneous velocity,

$$\mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{d}{dt} \left( v_r \mathbf{\hat{r}} + v_t \mathbf{\hat{\theta}} \right) = \frac{dv_r}{dt} \mathbf{\hat{r}} + v_r \frac{d\mathbf{\hat{r}}}{dt} + \frac{dv_t}{dt} \mathbf{\hat{\theta}} + v_t \frac{d\mathbf{\hat{\theta}}}{dt} \quad (8.43)$$

Noting that as it can be noted in the expressions, both unit vectors are time dependent and we have

$$\frac{d\mathbf{\hat{r}}}{dt} = \frac{d}{dt} \left( \cos(\theta) \mathbf{\hat{x}} + \sin(\theta) \mathbf{\hat{y}} \right) = \frac{d\hat{\theta}}{dt} \left( -\sin(\theta) \mathbf{\hat{x}} + \cos(\theta) \mathbf{\hat{y}} \right) = \frac{d\hat{\theta}}{dt} \mathbf{\hat{\theta}} = \omega \mathbf{\hat{\theta}} \quad (8.44)$$

and

$$\frac{d\hat{\theta}}{dt} = \frac{d}{dt} \left( -\sin(\theta) \mathbf{\hat{x}} + \cos(\theta) \mathbf{\hat{y}} \right) = \frac{d\hat{\theta}}{dt} \left( -\cos(\theta) \mathbf{\hat{x}} - \sin(\theta) \mathbf{\hat{y}} \right)$$

$$\Rightarrow \frac{d\hat{\theta}}{dt} = -\sin(\theta) \frac{d\hat{\theta}}{dt} = -\omega \mathbf{\hat{r}}. \quad (8.45)$$
so that the acceleration can be put in the form

\[
\vec{a}(t) = \frac{dv_r}{dt} \vec{r} + v_r \omega \dot{\theta} + \frac{dv_\theta}{dt} \vec{\theta} - v_\theta \omega \dot{r} = \left( \frac{dv_r}{dt} - v_\theta \omega \right) \vec{r} + \left( v_r \omega + \frac{dv_\theta}{dt} \right) \dot{\theta},
\]

(8.46)

In terms of the magnitude of the acceleration in the radial and angular directions

\[
a_r = \frac{dv_r}{dt} = \frac{d^2 r}{dt^2}
\]

(8.47)

and

\[
a_t = \frac{dv_\theta}{dt} = \frac{d(r \omega)}{dt} = \omega \frac{dr}{dt} + \frac{d\omega}{dt},
\]

(8.48)

the instantaneous acceleration becomes

\[
\vec{a}(t) = (a_r - v_\theta \omega) \hat{r} + (v_r \omega + a_t) \hat{\theta}.
\]

(8.49)

This can also be put in the form

\[
\vec{a}(t) = \left( \frac{d^2 r}{dt^2} - r \omega^2 \right) \hat{r} + \left( 2 \frac{d^2 \theta}{dt^2} - \frac{dr}{dt} \omega + r \alpha \right) \hat{\theta},
\]

(8.50)

where we used

\[
a_t = \omega \frac{dr}{dt} + \frac{d\omega}{dt}, a_r = \frac{d^2 r}{dt^2}, v_t = \omega r
\]

(8.51)

and the angular acceleration defined by

\[
\alpha = \frac{d^2 \theta}{dt^2}
\]

(8.52)

For a circular orbit where \( r \) is constant of motion (i.e. \( \vec{r} = R \hat{r} \)), we have

\[
\frac{dr}{dt} = 0, \quad \frac{d^2 r}{dt^2} = 0
\]

(8.53)

and

\[
\vec{v}(t) = v_r \hat{r} + v_\theta \hat{\theta} = \omega R \hat{\theta}
\]

(8.54)

so that

\[
\vec{a}(t) = -R \omega^2 \hat{r} + R \alpha \hat{\theta} = -\frac{v_r^2}{R} \hat{r} + R \omega \hat{\theta} = \vec{a}_c + \vec{a}_t
\]

(8.55)

where

\[
\vec{a}_c = -\frac{v_r^2}{R} \hat{r}
\]

(8.56)

and

\[
\vec{a}_t = R \alpha \hat{\theta} = R \frac{d\omega}{dt} \hat{\theta}
\]

(8.57)

are the centripetal and tangential accelerations, respectively.
8.2 Differential Operators

8.2.1 The gradient operator and directional derivative

In the previous section we considered the derivative of scalar and vector functions of one variable, time $t$. In most physical problems we often encounter products of scalars and vectors that are functions of two or more variables. Suppose we are given the scalar function in Cartesian coordinates, $f(x; y; z)$. We recall that the in\textsuperscript{finitesimal} change in this function, $df$ (the total differential), is given by

$$\text{df} = \frac{\partial f}{\partial x} \text{d}x + \frac{\partial f}{\partial y} \text{d}y + \frac{\partial f}{\partial z} \text{d}z.$$  \hfill (8.58)

In view of the expression for scalar vector product of two vectors $\vec{A}$ and $\vec{B}$ in Cartesian coordinates

$$\vec{A} \cdot \vec{B} = (A_x \hat{x} + A_y \hat{y} + A_z \hat{z}) \cdot (B_x \hat{x} + B_y \hat{y} + B_z \hat{z}) = A_x B_x + A_y B_y + A_z B_z,$$  \hfill (8.59)

one can re-write $df$ as

$$df = (dx \hat{x} + dy \hat{y} + dz \hat{z}) \cdot \left( \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z} \right).$$  \hfill (8.60)

Noting that the differential for the position vector in Cartesian coordinates, $\vec{r} = x \hat{x} + y \hat{y} + z \hat{z}$, is given by

$$d\vec{r} = dx \hat{x} + dy \hat{y} + dz \hat{z},$$  \hfill (8.61)

one can put $df$ in the form

$$df = d\vec{r} \cdot \left( \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z} \right) = d\vec{r} \cdot \left( \frac{\partial}{\partial x} \hat{x} \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z} \right) f,$$  \hfill (8.62)

where we factored out the function $f$ to the right side of the partial differential operators. The operators in the bracket are represented by the gradient operator,$\nabla$,

$$\nabla = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}.$$  \hfill (8.63)

The gradient operator operates on a scalar function, $f(x; y; z)$ to generate a vector function

$$\nabla f = \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z}.$$  \hfill (8.64)

Thus the in\textsuperscript{finitesimal} change, $df$, in the function $f(x; y; z)$ becomes

$$df = \nabla f \cdot d\vec{r}.$$  \hfill (8.65)
8.2. DIFFERENTIAL OPERATORS

**The directional derivative:** The function, $f(x, y, z)$, defines a surface in three dimensional space. Suppose this function is a constant, $f(x, y, z) = C$. Then the total differential for the function, $df$, becomes

$$df = \nabla f \cdot d\vec{r} = 0,$$

indicating that the scalar product of the vectors $\nabla f$ and $d\vec{r}$ is zero and the vectors are perpendicular to one another at the point $(x, y, z)$. Since $d\vec{r} = dx\hat{x} + dy\hat{y} + dz\hat{z}$, is an infinitesimal displacement on the surface defined by the function, $f(x, y, z)$, the vector $d\vec{r}$ is tangent to the surface at the point. Therefore, $\nabla f$, must be a vector normal to the surface at the same point.

Now let’s consider an infinitesimal displacement, $d\vec{s}$, on the surface defined by the function, $f(x, y, z)$. This displacement can be expressed as

$$d\vec{s} = ds\hat{n} = ds(n_x, n_y, n_z),$$

where

$$\hat{n} = n_x\hat{x} + n_y\hat{y} + n_z\hat{z}$$

is a unit vector along the direction of $d\vec{s}$. Thus one can write

$$d\vec{s} = ds\hat{n} = d\vec{r}; \Rightarrow ds(n_x, n_y, n_z) = (dx, dy, dz)$$

there follows that

$$dsn_x = dx, dsn_y = dy, dsn_z = dz \Rightarrow n_x = \frac{dx}{ds}, n_y = \frac{dy}{ds}, n_z = \frac{dz}{ds}.$$  

(8.69)

For a function $f(x, y, z)$, the directional derivative is the derivative along the direction $\hat{n}$. It is given by

$$\frac{df}{ds} = \frac{\partial f}{\partial x} \frac{dx}{ds} + \frac{\partial f}{\partial y} \frac{dy}{ds} + \frac{\partial f}{\partial z} \frac{dz}{ds}.$$

(8.70)

Using the results above, we may write

$$\frac{df}{ds} = \frac{\partial f}{\partial x} n_x + \frac{\partial f}{\partial y} n_y + \frac{\partial f}{\partial z} n_z = \left( \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z} \right) \cdot (n_x\hat{x} + n_y\hat{y} + n_z\hat{z})$$

(8.71)

and the directional derivative can then be expressed as

$$\frac{df}{ds} = \nabla f \cdot \hat{n},$$

where $\hat{n}$ is the unit vector along the direction of the vector $\vec{S}$ that is tangent to the surface defined by the function, $f(x, y, z)$ at a point on the surface with coordinates $(x, y, z)$. 

The Gradient Operator in curvilinear Coordinates: In the previous chapter we have introduced two curvilinear coordinate systems, namely Cylindrical and spherical coordinates system. Using the relation between the Cartesian coordinates and these curvilinear coordinates, it can be derived that the gradient in Cylindrical Coordinates is given by

$$\nabla f = \hat{r} \frac{\partial f}{\partial r} + \hat{\theta} \frac{\partial f}{\partial \theta} + \hat{z} \frac{\partial f}{\partial z}$$  \hspace{1cm} (8.72)$$

and in Spherical Coordinates by

$$\nabla f = \hat{r} \frac{\partial f}{\partial r} + \frac{\hat{\theta}}{r} \frac{\partial f}{\partial \theta} + \frac{\hat{\phi}}{r \sin(\theta)} \frac{\partial f}{\partial \phi}$$  \hspace{1cm} (8.73)$$

where \((\hat{r}, \hat{\theta}, \hat{z})\) and \((\hat{r}, \hat{\theta}, \hat{\phi})\) are the unit vectors in Cylindrical and spherical coordinates, respectively. Note that with the exception of \(\hat{z}\), these unit vectors are not constant unlike the Cartesian unit vectors.

Example 8.2 A point charge \(Q\) is located at the origin. The electrostatic potential at a point \(P\) a distance \(r\) from the point charge is given by

$$\phi (r) = \frac{kQ}{r},$$  \hspace{1cm} (8.74)$$

where \(k\) is the Coulomb’s law constant

$$k = \frac{1}{4 \pi e_0},$$  \hspace{1cm} (8.75)$$

where \(e_0\) is the electrical permittivity of a free space. In classical electrodynamics, the electric field is related to the electrostatic potential, \(\phi (r)\), by the equation

$$\vec{E} (\vec{r}) = -\nabla \phi (r).$$  \hspace{1cm} (8.76)$$

Find an expression for the electric field at point \(P\) using

(a)$$\nabla = \hat{r} \frac{\partial}{\partial r} + \hat{\theta} \frac{\partial}{r \partial \theta} + \frac{\hat{\phi}}{r \sin(\theta)} \frac{\partial}{\partial \phi}$$  \hspace{1cm} (8.77)$$

(b)$$\nabla = \frac{\partial}{\partial x} \hat{x} + \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z}$$  \hspace{1cm} (8.78)$$

(c) Find the directional derivative of the electrostatic potential at \((1, 2, -1)\) in the direction \(\vec{S} = 2\hat{x} - 2\hat{y} + \hat{z}\)

Solution:
8.2. DIFFERENTIAL OPERATORS

(a) Since the electric potential depends only on \( r \), using spherical coordinates is easier than Cartesian. Using the gradient in spherical coordinates, one can then write the electric field as

\[
\vec{E} \left( \vec{r} \right) = -\nabla \phi \left( r \right) = -\hat{r} \frac{\partial}{\partial r} \left( \frac{kQ}{r} \right) + \hat{\theta} \frac{1}{r} \frac{\partial}{\partial \theta} \left( \frac{kQ}{r} \right) + \hat{\varphi} \frac{1}{r \sin(\theta)} \frac{\partial}{\partial \varphi} \left( \frac{kQ}{r} \right)
\]

\[
\Rightarrow \vec{E} \left( \vec{r} \right) = -\hat{r} \frac{\partial}{\partial r} \left( \frac{kQ}{r} \right) = \frac{kQ}{r^2} \hat{r}
\]

(b) In Cartesian coordinates, the electric potential is given by

\[
\phi \left( r \right) = \frac{kQ}{r \left( x, y, z \right)},
\]

where

\[
\vec{r} = x \hat{x} + y \hat{y} + z \hat{z} \Rightarrow r = \sqrt{x^2 + y^2 + z^2}.
\]

Then we may write

\[
\vec{E} \left( \vec{r} \right) = -\nabla \phi \left( r \right) = - \left[ \frac{\partial \phi \left( r \right)}{\partial x} \hat{x} + \frac{\partial \phi \left( r \right)}{\partial y} \hat{y} + \frac{\partial \phi \left( r \right)}{\partial z} \hat{z} \right].
\]

Noting that the partial derivative, for example with respect to \( x \), can be determined as

\[
\frac{\partial \phi \left( r \right)}{\partial x} = \frac{d \phi \left( r \right)}{dr} \frac{dr}{dx} = \frac{d}{dr} \left[ \frac{kQ}{r} \right] \frac{dx}{\sqrt{x^2 + y^2 + z^2}} = -kQ \frac{x}{r^2} = -kQ \frac{x}{r^3},
\]

Similarly, the partial derivatives with respect to \( y \) and \( z \), we have

\[
\frac{\partial \phi \left( r \right)}{\partial y} = -kQ \frac{y}{r^2} \quad \frac{\partial \phi \left( r \right)}{\partial z} = -kQ \frac{z}{r^2}
\]

so that the gradient of \( \phi \left( \vec{r} \right) \) becomes

\[
\nabla \phi \left( \vec{r} \right) = -\frac{kQ}{r^3} x \hat{x} - \frac{kQ}{r^3} y \hat{y} - \frac{kQ}{r^3} z \hat{z} = -\frac{kQ}{r^3} \left( x \hat{x} + y \hat{y} + z \hat{z} \right) = -\frac{kQ}{r^3} \vec{r}.
\]

Noting that the unit vector along \( \vec{r} \) is

\[
\hat{r} = \frac{\vec{r}}{r}
\]

the electric field is found to be

\[
\vec{E} \left( \vec{r} \right) = -\nabla \phi \left( \vec{r} \right) = \frac{kQ}{r^2} \hat{r}.
\]
(c) We are given
\[ \vec{S} = 2\hat{x} - 2\hat{y} + \hat{z} \]  
and we want to find
\[ \frac{d\phi(r)}{ds} = \nabla \phi(r) \cdot \hat{n} \Big|_{(1,2,-1)}. \]  

Here \( \hat{n} \) is the unit vector along the vector \( \vec{S} \) tangent to the surface defined by the electrostatic potential function \( \phi(r) \) at a point \((x,y,z)\). This unit vector is given by
\[ \hat{n} = \frac{\vec{S}}{|\vec{S}|} = \frac{2}{3}\hat{x} - \frac{2}{3}\hat{y} + \frac{1}{3}\hat{z}. \]  

We are interested in the change in the potential, \( \phi(r) \), along this direction at a point, \((1,2,-1)\), on the surface. This is given by the directional derivative
\[ \frac{d\phi(r)}{ds} = \nabla \phi(r) \cdot \hat{n} \Big|_{(1,2,-1)}. \]  

Using the result we obtained in part (a) or (b) and the unit vector, \( \hat{n} \), one can then write
\[ \frac{d\phi(r)}{ds} = -\frac{kQ}{r^3} \left( x\hat{x} + y\hat{y} + z\hat{z} \right) \cdot \left( \frac{2}{3}\hat{x} - \frac{2}{3}\hat{y} + \frac{1}{3}\hat{z} \right) \Big|_{(1,2,-1)} \]  

Using the expression for \( r \) in Cartesian coordinates, we have
\[ \frac{d\phi(r)}{ds} = -\frac{kQ}{(x^2 + y^2 + z^2)^{3/2}} \left( \frac{2}{3}x - \frac{2}{3}y + \frac{1}{3}z \right) \Big|_{(1,2,-1)} \]  

so that the directional derivative for the electrostatic potential becomes
\[ \frac{d\phi(r)}{ds} = -\frac{kQ}{\sqrt{6}} \left( \frac{2}{3} - \frac{4}{3} + \frac{1}{3} \right) = \frac{kQ}{\sqrt{6}} \]  

**Exercise:** Do you thing we can find the angle between the electric field vector \( \vec{E}(\vec{r}) \) at point \((1,2,-1)\) (if we assume \( KQ = 1 \)) and \( \vec{S} \)?

### 8.2.2 The Divergence, the Curl, and the Laplacian

We have seen the gradient operator on a scalar function. It can provide information regarding how the scalar function changes along a given direction. Now we will consider how the gradient operator acts on a vector functions, such as the electric field due to a point charge we determined in the previous example,
8.2. DIFFERENTIAL OPERATORS

gravitational field due to a given mass, or magnetic field due to a current carrying wire. For vector functions the gradient operator acts in two different ways to describe the divergence and the curl (rotation) of the vector function. For a vector function

\[
\vec{V}(x, y, z) = V_x(x, y, z) \hat{x} + V_y(x, y, z) \hat{y} + V_z(x, y, z) \hat{z}
\]  

(8.96)

the divergence of \( \vec{V}(x, y, z) \), denoted by \( \nabla \cdot \vec{V} \) is given by

\[
\nabla \cdot \vec{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}.
\]  

(8.97)

Note that the divergence of a vector function results in a scalar function. For the curl of the vector, \( \vec{V}(x, y, z) \), is denoted by \( \nabla \times \vec{V} \). In Cartesian Coordinates, \( \nabla \times \vec{V} \) is given by the determinant expression

\[
\nabla \times \vec{V} = \begin{vmatrix}
\hat{x} & \hat{y} & \hat{z} \\
\frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\
V_x & V_y & V_z
\end{vmatrix}
\]  

(8.98)

that leads to

\[
\nabla \times \vec{V} = \left( \frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z} \right) \hat{x} + \left( \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x} \right) \hat{y} + \left( \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right) \hat{z}.
\]  

(8.99)

Suppose the vector field \( \vec{V}(x, y, z) \) is expressible as a gradient of some scalar function, \( \phi(x, y, z) \),

\[
\vec{V}(x, y, z) = \nabla \phi(x, y, z),
\]  

(8.100)

the divergence of this vector becomes

\[
\nabla \cdot \vec{V} = \nabla \cdot [\nabla \phi(x, y, z)] = (\nabla \cdot \nabla) \phi(x, y, z) = \nabla^2 \phi(x, y, z),
\]  

(8.101)

where

\[
\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}
\]  

(8.102)

is called the Laplacian operator. As an example, let’s consider the electric field vector, which is given by

\[
\vec{E}(\vec{r}) = -\nabla \phi(\vec{r}).
\]  

(8.103)

The divergence of \( \vec{E}(\vec{r}) \) leads to

\[
\nabla \cdot \vec{E}(\vec{r}) = -\nabla \cdot \nabla \phi(x, y, z) = -\left( \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} \right) = -\nabla^2 \phi(x, y, z).
\]  

(8.104)

In classical electrodynamics the divergence of the electric field is related to the volume charge density, \( \rho(x, y, z) \), by the differential form of Gauss’ law,

\[
\nabla \cdot \vec{E}(\vec{r}) = \frac{\rho(x, y, z)}{\varepsilon_0},
\]  

(8.105)
that can be expressed using the Laplacian as

\[ \nabla \cdot \vec{E}(\vec{r}) = -\nabla \cdot \nabla \phi(x, y, z) = - \left( \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} \right) = \frac{\rho(x, y, z)}{\varepsilon_0} \]

\[ \Rightarrow \nabla^2 \phi(x, y, z) = F(x, y, z). \tag{8.106} \]

where

\[ F(x, y, z) = -\frac{\rho(x, y, z)}{\varepsilon_0}. \tag{8.107} \]

Equation (8.106) is known as Poisson’s equation. In a free space where there is no charge, \( \rho(x, y, z) = 0 \), we find

\[ \nabla^2 \phi(x, y, z) = 0, \tag{8.108} \]

which is known as Laplace’s equation. Often you encounter Poisson’s and Laplace’s equations in classical electromagnetism and mechanics theories.

You also encounter, more frequently, in most branches of theoretical physics that could involve the gradient, the divergence, and the curl of a scalar and/or a vector functions. Thus it worth to remember the following relations:

(a) The divergence of a vector and a scalar: For a vector function, \( \vec{V} \), multiplied by a scalar functions \( f \), the divergence can be expressed as

\[ \nabla \cdot \left( f \vec{V} \right) = \vec{V} \cdot \nabla f + f \nabla \cdot \vec{V} \tag{8.109} \]

(b) The curl of the curl: In view of the relation for triple vector product

\[ \vec{A} \times \left( \vec{B} \times \vec{C} \right) = \vec{B} \left( \vec{A} \cdot \vec{C} \right) - \vec{C} \left( \vec{A} \cdot \vec{B} \right) \tag{8.110} \]

for the curl of the curl of the vector function, \( \vec{V} \), one can establish the relation

\[ \nabla \times \left( \nabla \times \vec{V} \right) = \nabla \left( \nabla \cdot \vec{V} \right) - \nabla^2 \vec{V}. \tag{8.111} \]

that involves the divergence and the Laplacian of the a vector function. Thus for the Laplacian of a vector function one can write

\[ \nabla^2 \vec{V} = \nabla \left( \nabla \cdot \vec{V} \right) - \nabla \times \left( \nabla \times \vec{V} \right) \]

(c) The curl of a vector and a scalar: For a vector function, \( \vec{V} \), multiplied by a scalar functions \( f \), the curl can be expressed as

\[ \nabla \times \left( f \vec{V} \right) = (\nabla f) \times \vec{V} + f \left( \nabla \times \vec{V} \right) = f \left( \nabla \times \vec{V} \right) - \vec{V} \times \nabla f \tag{8.112} \]

**Example 8.3** Consider two functions—a scalar function and a vector function—defined as follows:

\[ f(\vec{r}) = xyz, \quad \vec{V}(x, y, z) = (xy, yz, zx). \tag{8.113} \]

At the point \( (1, -1, 1) \), evaluate the following quantities:
8.2. DIFFERENTIAL OPERATORS

(a) $\nabla f$

(b) $\nabla \cdot \vec{V}$

(c) $\nabla \times \vec{V}$

(d) $\nabla^2 f$

(e) $\nabla^2 \vec{V}$

(f) $\nabla^2 \left( f\vec{V} \right)$

Solution:

(a) The gradient of the scalar function, $f$, given by

$$\nabla f = \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z}. \quad (8.114)$$

For $f(x,y,z) = xyz$,

$$\nabla f = \frac{\partial (xyz)}{\partial x} \hat{x} + \frac{\partial (xyz)}{\partial y} \hat{y} + \frac{\partial (xyz)}{\partial z} \hat{z} = yz\hat{x} + x\hat{y} + xy\hat{z}. \quad (8.115)$$

Upon evaluating this at $(x,y,z) = (1,-1,1)$, we find

$$\nabla f = \hat{x} + \hat{y} - \hat{z}. \quad (8.116)$$

(b) The divergence of the vector function, $\vec{V} = (V_x, V_y, V_z)$, is given by

$$\nabla \cdot \vec{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}. \quad (8.117)$$

Noting that $V_x = xy, V_y = yz, V_z = xz$

one finds

$$\nabla \cdot \vec{V} = \frac{\partial (xy)}{\partial x} + \frac{\partial (yz)}{\partial y} + \frac{\partial (xz)}{\partial z} = y + z + x. \quad (8.118)$$

At the point $(1,-1,1)$, the divergence is found to be

$$\nabla \cdot \vec{V} = 1. \quad (8.119)$$

(c) For $\nabla \times \vec{V}$, we have

$$\nabla \times \vec{V} = \det \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xy & yz & xz \end{vmatrix}. \quad (8.120)$$
that leads to
\[
\nabla \times \vec{V} = \left( \frac{\partial (xz)}{\partial y} - \frac{\partial (yz)}{\partial z} \right) \hat{x} + \left( \frac{\partial (xy)}{\partial z} - \frac{\partial (xz)}{\partial x} \right) \hat{y} \\
+ \left( \frac{\partial (yz)}{\partial x} - \frac{\partial (xy)}{\partial y} \right) \hat{z}
\]
\[
\Rightarrow \nabla \times \vec{V} = (0 - y) \hat{x} + (0 - z) \hat{y} + (0 - x) \hat{z} = -y\hat{x} - z\hat{y} - x\hat{z}. \quad (8.121)
\]
Evaluating this at the point \((1, -1, 1)\),
\[
\nabla \times \vec{V} = \hat{x} - \hat{y} - \hat{z}. \quad (8.122)
\]

(d) For the Laplacian of the scalar function, \(f\),
\[
\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \quad (8.123)
\]
we find
\[
\nabla^2 f = \frac{\partial^2 (xyz)}{\partial x^2} + \frac{\partial^2 (xyz)}{\partial y^2} + \frac{\partial^2 (xyz)}{\partial z^2} = 0. \quad (8.124)
\]
Noting that
\[
\nabla^2 f = \nabla \cdot \nabla f, \quad (8.125)
\]
one can also find the same result
\[
\nabla^2 f = \nabla \cdot (yz\hat{x} + xz\hat{y} + xy\hat{z}) = \frac{\partial (yz)}{\partial x} + \frac{\partial (xz)}{\partial y} + \frac{\partial (xy)}{\partial z} = 0, \quad (8.126)
\]
using the expression for the gradient of the function determined in (a).

(e) Here we want to determine the Laplacian of a vector function instead of a scalar function. In such cases, one should evaluate
\[
\nabla^2 \vec{V} = \frac{\partial^2 \vec{V}}{\partial x^2} + \frac{\partial^2 \vec{V}}{\partial y^2} + \frac{\partial^2 \vec{V}}{\partial z^2} \quad (8.127)
\]
where
\[
\vec{V} = V_x \hat{x} + V_y \hat{y} + V_z \hat{z}.
\]
Thus for
\[
\vec{V} = xy\hat{x} + yz\hat{y} + zx\hat{z}
\]
one finds
\[
\frac{\partial^2 \vec{V}}{\partial x^2} = \frac{\partial^2 (xy\hat{x} + yz\hat{y} + zx\hat{z})}{\partial x^2} = 0 \quad (8.128)
\]
\[
\frac{\partial^2 \vec{V}}{\partial y^2} = \frac{\partial^2 (xy\hat{x} + yz\hat{y} + zx\hat{z})}{\partial y^2} = 0 \quad (8.129)
\]
\[
\frac{\partial^2 \vec{V}}{\partial z^2} = \frac{\partial^2 (xy\hat{x} + yz\hat{y} + zx\hat{z})}{\partial z^2} = 0 \quad (8.130)
\]
The Laplacian becomes
\[ \nabla^2 \vec{V} = 0. \] (8.131)

We can also apply the relation
\[ \nabla^2 \vec{V} = \nabla \left( \nabla \cdot \vec{V} \right) - \nabla \times \left( \nabla \times \vec{V} \right) \]
along with the results we obtained earlier
\[ \nabla \cdot \vec{V} = y + z + x \] (8.132)
and
\[ \vec{U} = \nabla \times \vec{V} = -y \hat{x} - z \hat{y} - x \hat{z}. \] (8.133)

Noting that
\[ \nabla \left( \nabla \cdot \vec{V} \right) = \nabla (y + z + x) = \hat{x} \frac{\partial}{\partial x} (y + z + x) + \hat{y} \frac{\partial}{\partial y} (y + z + x) \]
\[ + \hat{z} \frac{\partial f}{\partial z} (y + z + x) \Rightarrow \nabla \left( \nabla \cdot \vec{V} \right) = \hat{x} + \hat{y} + \hat{z} \] (8.134)
and
\[ \nabla \times \left( \nabla \times \vec{V} \right) = \nabla \times \vec{U} = \det \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y & -z & -x \end{vmatrix} \] (8.135)
that can be simplified into
\[ \nabla \times \left( \nabla \times \vec{V} \right) = \left( \frac{\partial (y - z)}{\partial y} - \frac{\partial (-x)}{\partial z} \right) \hat{x} + \left( \frac{\partial (-y)}{\partial z} - \frac{\partial (-x)}{\partial x} \right) \hat{y} \]
\[ + \left( \frac{\partial (-z)}{\partial x} - \frac{\partial (-y)}{\partial y} \right) \hat{z} \Rightarrow \nabla \times \left( \nabla \times \vec{V} \right) = \hat{x} + \hat{y} + \hat{z}. \] (8.136)

Using these results, one also finds
\[ \nabla^2 \vec{V} = \nabla \left( \nabla \cdot \vec{V} \right) - \nabla \times \left( \nabla \times \vec{V} \right) = 0 \] (8.137)

(f) Noting that
\[ \nabla^2 \left( f \vec{V} \right) = \hat{x} \nabla^2 (f V_x) + \hat{y} \nabla^2 (f V_y) + \hat{z} \nabla^2 (f V_z) \] (8.138)
and
\[ \nabla^2 (f V_x) = \frac{\partial^2 ((xy)^2 z)}{\partial x^2} + \frac{\partial^2 ((xy)^2 z)}{\partial y^2} + \frac{\partial^2 ((xy)^2 z)}{\partial z^2} = 2y^2 z + 2x^2 z \] (8.139)
\[ \nabla^2 (fV_y) = \frac{\partial^2 ((yz)^2 x)}{\partial x^2} + \frac{\partial^2 ((yz)^2 x)}{\partial y^2} + \frac{\partial^2 ((yz)^2 x)}{\partial z^2} = 2z^2x + 2y^2x \] (8.140)

\[ \nabla^2 (fV_z) = \frac{\partial^2 ((zx)^2 y)}{\partial x^2} + \frac{\partial^2 ((zx)^2 y)}{\partial y^2} + \frac{\partial^2 ((zx)^2 y)}{\partial z^2} = 2z^2y + 2x^2y \] (8.141)

we find

\[ \nabla^2 \left( f \hat{\nabla} \right) = \hat{x} (2y^2z + 2x^2z) + \hat{y} (2z^2x + 2y^2x) + \hat{z} (2z^2y + 2x^2y) \] (8.142)

Using Mathematica

\[ \text{Out[1]} = \langle \langle \text{Calculus`VectorAnalysis`} \rangle \rangle \]

\[ \text{Out[2]} = \text{Grad}[[x \, y \, z, \text{Cartesian}[x, y, z]]] \]

\[ \text{Out[3]} = (y \, z, \hat{x} \, z, \hat{y}) \]

\[ \text{Out[4]} = \text{Div}[[x \, y, z \, z, z \times \{x, y, z\}, \text{Cartesian}[x, y, z]]] \]

\[ \text{Out[5]} = x \, y + z \]

\[ \text{Out[6]} = \text{Curl}[[x \, y, z \, z, z \times \{x, y, z\}, \text{Cartesian}[x, y, z]]] \]

\[ \text{Out[7]} = (-y, -z, -x) \]

\[ \text{Out[8]} = \text{Laplacian}[[x \, y \, z, \text{Cartesian}[x, y, z]]] \]

\[ \text{Out[9]} = 0 \]

The gradient, the divergence, the Laplacian, and the curl that we have seen so far is for scalar and vector functions expressed in terms of Cartesian coordinates. Depending on the problem we want to solve, sometimes it is convenient to evaluate these quantities in curvilinear coordinates by expressing the functions using the corresponding transformations for the Cartesian to the convenient curvilinear coordinates. Thus it is important to know the expression for these quantities in the two curvilinear coordinates we studied (cylindrical and spherical coordinates). These expressions can be derived from its corresponding expressions in Cartesian. For now we list these expressions in cylindrical and spherical coordinates.

(a) **Cylindrical coordinates** \((r, \varphi, z)\): For a scalar function, \(f(r, \varphi, z)\), the gradient and Laplacian are given by

\[ \nabla f = \hat{r} \frac{\partial f}{\partial r} + \hat{\varphi} \frac{1}{r} \frac{\partial f}{\partial \varphi} + \hat{z} \frac{\partial f}{\partial z} \] (8.143)
8.2. DIFFERENTIAL OPERATORS

and
\[ \nabla^2 f = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}, \] (8.144)

respectively. For a vector function,
\[ \vec{V}(r, \varphi, z) = V_r (r, \varphi, z) \hat{r} + V_\varphi (r, \varphi, z) \hat{\varphi} + V_z (r, \varphi, z) \hat{z}, \]
the divergence and the curl can determined using
\[ \nabla \cdot \vec{V} = \frac{1}{r} \frac{\partial}{\partial r} (rV_r) + \frac{1}{r} \frac{\partial}{\partial \varphi} (V_\varphi) + \frac{\partial}{\partial z} (V_z) \] (8.145)

and
\[ \nabla \times \vec{V} = \hat{r} \left( \frac{1}{r} \frac{\partial V_z}{\partial \varphi} - \frac{\partial V_\varphi}{\partial z} \right) + \hat{\varphi} \left( \frac{\partial V_z}{\partial r} - \frac{\partial V_r}{\partial \varphi} \right) + \hat{z} \left( \frac{\partial}{\partial r} (rV_r) - \frac{\partial V_\varphi}{\partial \varphi} \right), \] (8.146)
respectively.

(b) Spherical coordinates \((r, \theta, \varphi)\): For a scalar function, \(f(r, \theta, \varphi)\), the gradient can be expressed as
\[ \nabla f = \hat{r} \frac{\partial f}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial f}{\partial \theta} + \hat{\varphi} \frac{1}{r \sin(\theta)} \frac{\partial f}{\partial \varphi}, \] (8.147)

and the Laplacian as
\[ \nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left( \sin(\theta) \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2(\theta)} \frac{\partial^2 f}{\partial \varphi^2}. \] (8.148)

On the other hand for a vector function,
\[ \vec{V}(r, \varphi, \theta) = V_r (r, \varphi, \theta) \hat{r} + V_\theta (r, \varphi, \theta) \hat{\theta} + V_\varphi (r, \varphi, \theta) \hat{\varphi}, \]
the divergence and the curl are given by
\[ \nabla \cdot \vec{V} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 V_r \right) + \frac{1}{r \sin(\theta)} \frac{\partial}{\partial \theta} \left( \sin(\theta) V_\theta \right) + \frac{1}{r \sin(\theta)} \frac{\partial V_\varphi}{\partial \varphi}, \] (8.149)

and
\[ \nabla \times \vec{V} = \hat{r} \left( \frac{\partial}{\partial \theta} (\sin(\theta) V_\phi) \right) + \hat{\theta} \left( \frac{1}{r \sin(\theta)} \frac{\partial V_\phi}{\partial \varphi} - \frac{1}{r} \frac{\partial (rV_\phi)}{\partial r} \right) + \hat{\varphi} \frac{1}{r} \left( \frac{\partial}{\partial r} (rV_\theta) \right) \] (8.150)
respectively.
Example 8.4 Consider an infinitely long cylindrical wire of radius $a$ carrying a current $I$ along the positive $z$-direction and this current is uniformly distributed.

The magnetic field is given by

$$\vec{B} = \left( \mu_0 I r / 2 \pi a^2 \right) \hat{\phi} \text{ for } r < a$$

(8.151)

at a distance $r$ from the wire in cylindrical coordinates $(r, \phi, z)$, $\mu_0$ is magnetic permeability of a free space.

(a) Using parametric plot show that the magnetic field lines form a concentric circle.

(b) Show that magnetic field vector is a rotational vector field (i.e. the $\nabla \times \vec{B} \neq 0$) inside the cylinder.

(c) Show that $\nabla \cdot \vec{B} = 0$.

Solution:

(a)
We note that the components of the magnetic field in cylindrical coordinates are

\[ B_r = 0, B_\varphi = \frac{\mu_0 I_r}{2\pi a^2}, B_z = 0. \]

Then using the expression for the curl of a vector in cylindrical coordinates

\[ \nabla \times \vec{B} = \hat{r} \left( \frac{1}{r} \frac{\partial B_\varphi}{\partial r} - \frac{\partial B_z}{\partial \varphi} \right) + \hat{\varphi} \left( \frac{\partial B_r}{\partial z} - \frac{\partial B_z}{\partial r} \right) + \hat{z} \left( \frac{1}{r} \frac{\partial}{\partial r} (r B_\varphi) - \frac{\partial B_r}{\partial \varphi} \right) \]

we find

\begin{align*}
\nabla \times \vec{B} &= -\hat{r} \left( \frac{1}{r} \frac{\partial B_\varphi}{\partial r} \right) + \hat{\varphi} \left( \frac{\partial B_r}{\partial z} - \frac{\partial B_z}{\partial r} \right) + \hat{z} \left( \frac{1}{r} \frac{\partial}{\partial r} (r B_\varphi) \right) \\
&= -\hat{r} \frac{\partial}{\partial z} \left[ \frac{\mu_0 I_r}{2\pi a^2} \right] + \hat{\varphi} \frac{1}{r} \frac{\partial}{\partial r} \left( \frac{\mu_0 I_r}{2\pi a^2} \right) \\
&= \nabla \times \vec{B} = \hat{z} \left[ \frac{1}{r} \frac{\partial}{\partial r} \left( \frac{\mu_0 I_r^2}{2\pi a^2} \right) \right] = \frac{\mu_0 I}{\pi a^2} \hat{z}.
\end{align*}
CHAPTER 8. VECTOR CALCULUS

Noting the current density, which is defined as
\[ J = \frac{\text{Current}}{\text{Area}}, \]  
(8.155)
is
\[ J = \frac{I}{\pi a^2} \]  
(8.156)
we may write
\[ \nabla \times \vec{B} = \mu_0 J \hat{z} = \mu_0 \vec{J} \]  
(8.157)
which is known as Ampere’s law.

(c) The divergence in cylindrical coordinates is given by
\[ \nabla \cdot \vec{B} = \frac{1}{r} \frac{\partial}{\partial r} (rV_r) + \frac{1}{r} \frac{\partial}{\partial \phi} (V_\phi) + \frac{\partial V_z}{\partial z} \]  
and using
\[ B_r = 0, B_\phi = \frac{\mu_0 I_r}{2\pi a^2}, B_z = 0. \]
we find
\[ \nabla \cdot \vec{B} = 0 \]

[Electricity and Magnetism I...PHYS 4310]

Example 8.5 In example 8.2 we have shown that the electric field of a point charge, \( Q \), position at the origin, is given
\[ \vec{E}(\vec{r}) = -\nabla \phi(\vec{r}) = \frac{kQ}{r^2} \hat{r}. \]  
(8.158)
Show that this electric field vector is irrotational vector field.

Solution:
The electric field vector in spherical coordinates has components
\[ E_r = \frac{kQ}{r^2}, E_\theta = E_\phi = 0 \]
then the curl of the electric field vector in spherical coordinates
\[ \nabla \times \vec{E} = \hat{r} \frac{1}{r \sin(\theta)} \left[ \frac{\partial}{\partial \theta} (\sin(\theta) E_\phi) - \frac{\partial E_\theta}{\partial \phi} \right] \]  
(8.159)
\[ + \hat{\theta} \left[ \frac{1}{r \sin(\theta)} \frac{\partial E_r}{\partial \phi} - \frac{1}{r} \frac{\partial}{\partial r} (rE_\phi) \right] + \hat{\phi} \left[ \frac{\partial}{\partial r} (rE_\theta) - \frac{\partial E_r}{\partial \theta} \right] \]  
(8.160)
becomes
\[ \nabla \times \vec{E} = 0. \]  
(8.161)
This means that the electric field vector is irrotational vector field.

Exercise: Find the divergence of the electric field and try to make a physical interpretation for the result.
8.3 Single vector function integrals

8.3.1 Line vector integrals

Example 8.6 Consider the vector field

\[ \vec{A}(\vec{r}) = x^2 y \hat{x} + xy^2 \hat{y} + a^3 e^{-\beta y} x \cos(\alpha) \hat{z}, \]  

(8.162)

where \( a, \alpha, \) and \( \beta \) are known constants. Evaluate the line integral

\[ I = \int_C \vec{A}(\vec{r}) \cdot d\vec{r} \]  

(8.163)

around the curve \( C \) shown in the diagram below.

Solution:

The closed line integral can easily be carried out if we divide the integral into three line integrals corresponding to the three sides \( s_1, s_2, \) and \( s_3, \)

\[ \int_C \vec{A}(\vec{r}) \cdot d\vec{r} = \int_{s_1} \vec{A}(\vec{r}) \cdot d\vec{r} + \int_{s_2} \vec{A}(\vec{r}) \cdot d\vec{r} + \int_{s_3} \vec{A}(\vec{r}) \cdot d\vec{r} \]  

(8.164)

Let’s represent a point on any of these line segments by a vector

\[ \vec{r} = x\hat{x} + y\hat{y} \]  

(8.165)

so that

\[ d\vec{r} = dx\hat{x} + dy\hat{y} \]  

(8.166)
For a point on side $s_1$

$$x = 0 \Rightarrow \vec{r} = y\hat{y} \Rightarrow d\vec{r} = dy\hat{y}$$  \hspace{1cm} (8.167)

for $s_2$

$$y = \sqrt{2} \Rightarrow \vec{r} = x\hat{x} + \sqrt{2}\hat{y} \Rightarrow d\vec{r} = dx\hat{x}$$  \hspace{1cm} (8.168)

and for $s_3$

$$y = \sqrt{x} \Rightarrow \vec{r} = x\hat{x} + \sqrt{x}\hat{y} \Rightarrow d\vec{r} = dx\hat{x} + \frac{dx}{2\sqrt{x}}\hat{y}.$$  \hspace{1cm} (8.169)

The function

$$\vec{A}(\vec{r}) = x^2 y\hat{y} + xy^2\hat{y} + a^3 e^{-\beta y} x \cos(\alpha) \hat{z},$$  \hspace{1cm} (8.170)

also takes different expressions on these three sides of the curve. On $s_1$

$$x = 0 \Rightarrow \vec{A}(\vec{r}) = 0$$  \hspace{1cm} (8.171)

on $s_2$

$$y = \sqrt{2}, z = 0 \Rightarrow \vec{A}(\vec{r}) = x^2 \sqrt{2}\hat{x} + 2x\hat{y} + a^3 e^{-\sqrt{2}\beta} x \cos(\alpha) \hat{z}$$  \hspace{1cm} (8.172)

and on $s_3$

$$y = \sqrt{x}, z = 0 \Rightarrow \vec{A}(\vec{r}) = x^{5/2}\hat{x} + x^2\hat{y} + a^3 e^{-\sqrt{x}\beta} x \cos(\alpha) \hat{z}$$  \hspace{1cm} (8.173)

Therefore, using the three line integrals on $s_1$

$$\int_{s_1} \vec{A}(\vec{r}) \cdot d\vec{r} = 0$$  \hspace{1cm} (8.174)

on $s_2$

$$\int_{s_2} \vec{A}(\vec{r}) \cdot d\vec{r} = \int_0^2 \left( x^2 \sqrt{2} \hat{x} + 2x\hat{y} \right) \cdot dx\hat{x} = \sqrt{2} \int_0^2 x^2 dx = \frac{\sqrt{2}}{3} x^3 \bigg|_0^2 = \frac{8\sqrt{2}}{3}$$  \hspace{1cm} (8.175)

and on $s_3$

$$\int_{s_3} \vec{A}(\vec{r}) \cdot d\vec{r} = \int_0^0 \left( x^{5/2}\hat{x} + x^2\hat{y} \right) \left( dx\hat{x} + \frac{dx}{2\sqrt{x}}\hat{y} \right) = \int_2^0 \left[ x^{5/2} + \frac{x^2}{2\sqrt{x}} \right] dx$$

$$= \frac{2}{5} x^{7/2} + \frac{1}{5} x^{3/2} \bigg|_2^0$$

$$= \int_{s_3} \vec{A}(\vec{r}) \cdot d\vec{r} = -\frac{16\sqrt{2}}{7} - \frac{4\sqrt{2}}{5}.$$  \hspace{1cm} (8.176)

we find

$$\oint_{C} \vec{A}(\vec{r}) \cdot d\vec{r} = \frac{8\sqrt{2}}{3} - \frac{16\sqrt{2}}{7} - \frac{4\sqrt{2}}{5} = -\frac{44\sqrt{2}}{105}$$  \hspace{1cm} (8.177)
8.3.2 Conservative vector fields

Conservative Vector Fields: a vector field \( \vec{V}(x, y, z) = V_x(x, y, z) \hat{x} + V_y(x, y, z) \hat{y} + V_z(x, y, z) \hat{z} \) (8.178)
is said to be conservative if any one of the following conditions is satisfied:

- If the line integral over a closed curve is zero independent of the path,
  \[ \oint_C \vec{V}(x, y, z) \cdot d\vec{r} = 0. \] (8.179)

- If the curl of the vector field is zero,
  \[ \nabla \times \vec{V}(x, y, z) = 0. \] (8.180)

- If the vector field can be expressed as a gradient of some scalar function \( U(x, y, z) \),
  \[ \vec{V}(x, y, z) = \nabla U(x, y, z). \] (8.181)

Gravitational and electric fields are conservative fields.

Example 8.7 Consider the moon and the earth. Imagine the origin of a spherical coordinate system coincides with the center of the earth. Then the position of the moon can be described by \((r, \theta, \varphi)\) in spherical coordinates. The gravitational potential energy of the earth and the moon is given by

\[ U = \frac{GMm}{r} \] (8.182)

where \( r \) is the distance between the earth and the moon, \( M \) is mass of the earth, and \( m \) is mass of the moon. (See Fig. 8.1)

Figure 8.1: Moon’s orbit around the earth.
(a) Find the gravitational force

\[ \mathbf{F} = -\nabla \cdot U \]  \hspace{1cm} (8.183)

on the moon using the gravitational potential energy.

(b) Show that gravitational force is a conservative force using the work done by the gravitational force when the moon makes one complete revolution around the earth.

\[ W = \int_{C} \mathbf{F} \cdot d\mathbf{r} \]

Solution:

(a) The gravitational force is given by

\[ \mathbf{F} = -\nabla \cdot U \]  \hspace{1cm} (8.184)

using the gradient in spherical coordinates

\[ \nabla U = \hat{r} \frac{\partial U}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial U}{\partial \theta} + \hat{\varphi} \frac{1}{r \sin(\theta)} \frac{\partial U}{\partial \varphi} \]  \hspace{1cm} (8.185)

we find

\[ \mathbf{F} = -\hat{r} \frac{\partial U}{\partial r} = -\hat{r} \frac{\partial}{\partial r} \left[ \frac{GMm}{r} \right] = \frac{GMm}{r^2} \hat{r} \]  \hspace{1cm} (8.186)

(b) The work done is given by

\[ W = \int_{C} \mathbf{F} \cdot d\mathbf{r} = \int_{C} \frac{GMm}{a^2} \hat{r} \cdot d\mathbf{r} \]  \hspace{1cm} (8.187)

for a circular orbit \( r \) is a constant. That means in spherical coordinates the position vector

\[ \mathbf{r} = a \hat{r} \]  \hspace{1cm} (8.188)

becomes

\[ d\mathbf{r} = ad\hat{r}. \]  \hspace{1cm} (8.189)

Recalling that in spherical coordinates

\[ \mathbf{r} = r \sin(\theta) \cos(\varphi) \hat{x} + r \sin(\theta) \sin(\varphi) \hat{y} + r \cos(\theta) \hat{z} \]

the unit vector along the radial direction

\[ \hat{r} = \frac{\partial \mathbf{r}}{\partial r} \bigg/ \left| \frac{\partial \mathbf{r}}{\partial r} \right| \]

\[ = \sin(\theta) \cos(\varphi) \hat{x} + \sin(\theta) \sin(\varphi) \hat{y} + \cos(\theta) \hat{z} \]  \hspace{1cm} (8.190)

we may write

\[ d\mathbf{r} = [\cos(\theta) \cos(\varphi) d\theta - \sin(\theta) \sin(\varphi) d\varphi] \hat{x} \]
\[ + [\cos(\theta) \sin(\varphi) d\theta + \sin(\theta) \cos(\varphi) d\varphi] \hat{y} - \sin(\theta) d\theta \hat{z}. \]  \hspace{1cm} (8.191)
8.3. SINGLE VECTOR FUNCTION INTEGRALS

There follows that

\[ \hat{r} \cdot d\vec{r} = \sin(\theta) \cos(\theta) \cos^2(\varphi) \, d\theta - \sin^2(\theta) \sin(\varphi) \cos(\varphi) \, d\varphi + \sin(\theta) \cos(\theta) \sin^2(\varphi) \, d\theta + \sin^2(\theta) \sin(\varphi) \cos(\varphi) \, d\varphi - \sin(\theta) \cos(\theta) \, d\theta, \]

\[ \Rightarrow \hat{r} \cdot d\vec{r} = 0 \tag{8.192} \]

Then the work done

\[ W = \oint_C \frac{GMm}{a^2} \, \hat{r} \cdot d\vec{r} = 0 \tag{8.193} \]

8.3.3 Conservative Fields and Exact Differentials

Exact vs Inexact Differentials: Consider the scalar function \( U(x, y, z) \). We recall that the differential for this function is given by

\[ dU = \frac{\partial U}{\partial x} \, dx + \frac{\partial U}{\partial y} \, dy + \frac{\partial U}{\partial z} \, dz \]

\[ dU = \left( \frac{\partial U}{\partial x} \hat{x} + \frac{\partial U}{\partial y} \hat{y} + \frac{\partial U}{\partial z} \hat{z} \right) \cdot (dx \hat{x} + dy \hat{y} + dz \hat{z}) = \nabla U \cdot d\vec{r}. \tag{8.194} \]

Let’s assume that the scalar potential \( U(x, y, z) \) be a potential function for some conservative vector field, \( \vec{V}(x, y, z) = V_x(x, y, z) \hat{x} + V_y(x, y, z) \hat{y} + V_z(x, y, z) \hat{z} \) (i.e. \( \vec{V}(x, y, z) = \nabla U(x, y, z) \)) then we can write

\[ dU = \nabla U \cdot d\vec{r} = \vec{V} \cdot d\vec{r} \Rightarrow V_x = \frac{\partial U}{\partial x}, V_y = \frac{\partial U}{\partial y}, V_z = \frac{\partial U}{\partial z}. \tag{8.195} \]

For a conservative vector field, we know that

\[ \nabla \times \vec{V}(x, y, z) = \left| \begin{array}{ccc} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ V_x & V_y & V_z \end{array} \right| = \left( \frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z} \right) \hat{x} + \left( \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x} \right) \hat{y} + \left( \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right) \hat{z} = 0 \tag{8.196} \]

which leads to

\[ \frac{\partial V_z}{\partial y} = \frac{\partial V_y}{\partial z}, \frac{\partial V_x}{\partial z} = \frac{\partial V_z}{\partial x}, \frac{\partial V_y}{\partial x} = \frac{\partial V_z}{\partial y}. \tag{8.197} \]

Recalling that we assumed the scalar function \( U \) is a function related to a conservative vector field \( V \) by

\[ V_x = \frac{\partial U}{\partial x}, V_y = \frac{\partial U}{\partial y}, V_z = \frac{\partial U}{\partial z}. \tag{8.198} \]
we may write
\[ \frac{\partial V_z}{\partial y} = \frac{\partial V_y}{\partial z} \quad \text{or} \quad \frac{\partial^2 U}{\partial y \partial z} = \frac{\partial^2 U}{\partial z \partial y}, \]
\[ \frac{\partial V_y}{\partial x} = \frac{\partial V_x}{\partial y} \quad \text{or} \quad \frac{\partial^2 U}{\partial x \partial y} = \frac{\partial^2 U}{\partial y \partial x}. \quad (8.199) \]

These relations are known as the Reciprocity Relations. The differential
\[ dU = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy + \frac{\partial U}{\partial z} dz, \quad (8.200) \]
is said to be an exact differential if the Reciprocity Relations are fully satisfied. If not, then the differential is inexact. We derived the Reciprocity Relations assuming the scalar function \( U \) is a function that its gradient gives a conservative vector field \( \vec{V} \) (i.e. \( \vec{V}(x,y,z) = \nabla U(x,y,z) \)). Therefore, we can conclude that a differential
\[ dU = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy + \frac{\partial U}{\partial z} dz = f(x,y,z)dx + g(x,y,z)dy + h(x,y,z)dz \quad (8.201) \]
is an exact differential if the vector field constructed as
\[ \vec{V}(x,y,z) = f(x,y,z)\hat{x} + g(x,y,z)\hat{y} + h(x,y,z)\hat{z} \quad (8.202) \]
is a conservative vector field. Which means we can also say if
\[ \oint_C \vec{V}(x,y,z) \cdot d\vec{r} = \int_C f(x,y,z)dx + \int_C g(x,y,z)dy + \int_C h(x,y,z)dz = 0, \quad (8.203) \]
the differential is exact otherwise it is inexact.

## 8.4 Multiple vector integrals

### 8.4.1 Green’s Theorem

Let’s consider two functions \( P(x,y) \) and \( Q(x,y) \). If these two functions and its first derivatives are continuous for all \( x \) and \( y \) in the region bounded by the curve \( C \) of area \( A \), then Green’s theorem states
\[ \iint_A \left[ \frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y} \right] dxdy = \oint_C [P(x,y)dx + Q(x,y)dy], \quad (8.204) \]
where the line integral is counterclockwise around the boundary of the area \( A \) (i.e. on the curve \( C \)).

**Proof:**
We note that

\[ \iint_A \left[ \frac{\partial Q(x, y)}{\partial x} - \frac{\partial P(x, y)}{\partial y} \right] dxdy = \iint_A \frac{\partial Q(x, y)}{\partial x} dxdy - \iint_A \frac{\partial P(x, y)}{\partial y} dydx. \]  

(8.205)

For the area shown in the figure below we may write

\[ \iint_A \frac{\partial P(x, y)}{\partial y} dydx = \int_a^b \left[ \int_{y_l}^{y_u} \frac{\partial P(x, y)}{\partial y} dy \right] dx \]  

(8.206)

so that using the relation

\[ \int_a^b \frac{df(t)}{dt} dt = f(b) - f(a) \]  

(8.207)

we find

\[ \iint_A \frac{\partial P(x, y)}{\partial y} dydx = \int_a^b \left[ P(x, y_u) - P(x, y_l) \right] dx \]

\[ = \int_a^b P(x, y_u) dx - \int_a^b P(x, y_l) dx = - \int_b^a P(x, y_u) dx - \int_a^b P(x, y_l) dx \]

\[ \Rightarrow - \iint_A \frac{\partial P(x, y)}{\partial y} dydx = \oint_C P(x, y) dx. \]  

(8.208)

Note that the line integral over the closed curve \( C \) is the boundary of the area \( A \) and it must be integrated in a counterclockwise direction. Because the limits of integration in the two line integrals above indicate that we perform the integration first on the upper part of the curve from \( b \rightarrow a \) and then from \( a \rightarrow b \). Similarly if we perform the integration in the reverse
order (i.e. first with respect to \(x\) and then with respect to \(y\)) for the integral

\[
\iint_A \frac{\partial Q(x, y)}{\partial x} \, dx \, dy = \int_c^d \left[ \int_{x_i}^{x_f} \frac{\partial Q(x, y)}{\partial x} \, dx \right] \, dy \tag{8.209}
\]

as shown in the figure below, for the same area then we can write

\[
\iint_A \frac{\partial Q(x, y)}{\partial x} \, dx \, dy = \int_c^d [Q(x_r, y) - Q(x_l, y)] \, dy = \int_c^d Q(x_r, y) \, dy - \int_c^d Q(x_l, y) \, dy
\]

\[
= \int_c^d Q(x_r, y) \, dy + \int_d^c Q(x_l, y) \, dy \Rightarrow \iint_A \frac{\partial Q(x, y)}{\partial x} \, dx \, dy = \int_C Q(x, y) \, dy \tag{8.210}
\]

Here also the line integral over the same curve \(C\) is performed in the counterclockwise direction like the previous integration. Therefore, combining the two relations

\[
- \iint_A \frac{\partial P(x, y)}{\partial y} \, dy \, dx = \oint_C P(x, y) \, dx, \quad \iint_A \frac{\partial Q(x, y)}{\partial x} \, dx \, dy = \oint_C Q(x, y) \, dy, \tag{8.211}
\]

we find

\[
\iint_A \left[ \frac{\partial Q(x, y)}{\partial x} - \frac{\partial P(x, y)}{\partial y} \right] \, dx \, dy = \oint_C [P(x, y) \, dx + Q(x, y) \, dy]. \tag{8.212}
\]

**Example 8.8** Let the area \(A\) be the region bounded by the curves \(y = x^2\) and \(x = y^2\). Verify Green’s theorem for the integral

\[
I = \oint_C \left[ (2xy - x^2) \, dx + (x + y^2) \, dy \right] \tag{8.213}
\]
8.4. MULTIPLE VECTOR INTEGRALS

Solution:

Comparing the given integral with
\[
\iint_A \left[ \frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y} \right] \, dx \, dy = \oint_C \left[ P(x,y) \, dx + Q(x,y) \, dy \right]. \tag{8.214}
\]
we want to verify that the result of
\[
I_1 = \oint_C \left[ P(x,y) \, dx + Q(x,y) \, dy \right] = \oint_C \left[ (2xy - x^2) \, dx + (x + y^2) \, dy \right]
\]
is equal to
\[
I_2 = \iint_A \left[ \frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y} \right] \, dx \, dy
\]
\[
= \iint_A \left[ \frac{\partial (x + y^2)}{\partial x} - \frac{\partial (2xy - x^2)}{\partial y} \right] \, dx \, dy \tag{8.216}
\]
for the area shown in the figure below:

Another short cut for the line integral
\[
I_1 = \int_{C_1} \left[ (2xy - x^2) \, dx + (x + y^2) \, dy \right]
\]
\[
+ \int_{C_2} \left[ (2xy - x^2) \, dx + (x + y^2) \, dy \right]
\]
For the lower curve \( y = x^2 \Rightarrow dy = 2xdx \) and for the upper curve \( x = y^2 \Rightarrow dx = 2ydy \) and noting that the two curves intersect at two points \((0,0)\) and \((1,1)\), we can write

\[
I_1 = \int_0^1 [(2x^3 - x^2) dx + (x + x^4) 2xdx] + \int_1^0 [(2y^3 - y^4) 2ydy + 2y^2dy] \quad (8.217)
\]

\[
I_1 = \int_0^1 [2x^3 - x^2 + 2x^2 + 2x^5] dx + \int_1^0 [4y^4 - 2y^5 + 2y^2] dy \quad (8.218)
\]

\[
I_1 = \int_0^1 [2x^5 + 2x^3 + x^2] dx + \int_1^0 [4y^4 - 2y^5 + 2y^2] dy \quad (8.219)
\]

\[
I_1 = \left[ \frac{x^6}{3} + \frac{x^4}{2} + \frac{x^3}{3} \right]_0^1 + \left[ \frac{4y^5}{5} - \frac{y^6}{3} + \frac{2y^3}{3} \right]_0^1 = \frac{1}{30} \quad (8.220)
\]

For the area integral

\[
I_2 = \iint_A \left[ \frac{\partial (x + y^2)}{\partial x} - \frac{\partial (2xy - x^2)}{\partial y} \right] dxdy = \int_0^1 \int_{y^2}^{\sqrt{y}} (1 - 2x) dxdy
\]

\[
= \int_0^1 (x - x^2)|_{y^2}^{\sqrt{y}} dy = \int_0^1 (\sqrt{y} - y - y^2 + y^4) dy
\]

\[
= \left[ \frac{2}{3}y^{3/2} - \frac{y^2}{2} - \frac{y^3}{3} + \frac{y^5}{5} \right]_0^1 = \frac{2}{3} - \frac{1}{2} - \frac{1}{3} + \frac{1}{5} = \frac{1}{30} \quad (8.221)
\]

### 8.4.2 The Stokes’ Theorem

**Stokes’ Theorem:**

\[
\oint_{\text{Curve bounding } A} \vec{D} \cdot d\vec{r} = \int_{\text{Surface } A} \nabla \times \vec{D} \cdot \hat{n}dA. \quad (8.222)
\]

Consider the surface bounded by the curve, \( C \), on the x-y plane as shown in the figure below. Using an infinitesimal area \( dA \) in this region one can establish the relation

\[
\hat{n}dA = \hat{z}dxdy, \quad (8.223)
\]

where \( \hat{n} = \hat{z} \) is the unit vector normal to this area. On the curve, an infinitesimal displacement can be written as

\[
d\vec{r} = dx\hat{x} + dy\hat{y}. \quad (8.224)
\]

For a vector \( \vec{D} \) that depends on the \( x \) and \( y \) coordinates

\[
\vec{D}(x,y) = D_x(x,y)\hat{x} + D_y(x,y)\hat{y} + D_z(x,y)\hat{z} \quad (8.225)
\]
the curl in Cartesian coordinates is given by
\[
\nabla \times \vec{D} = \det \begin{vmatrix}
\frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\
\hat{x} & \hat{y} & \hat{z} \\
D_x & D_y & D_z
\end{vmatrix}
= \left( \frac{\partial D_z}{\partial y} - \frac{\partial D_y}{\partial z} \right) \hat{x} + \left( \frac{\partial D_z}{\partial x} - \frac{\partial D_x}{\partial z} \right) \hat{y} + \left( \frac{\partial D_y}{\partial x} - \frac{\partial D_x}{\partial y} \right) \hat{z}
\]
(8.226)

One can then write
\[
\vec{D} \cdot d\vec{r} = D_x(x, y) \, dx + D_y(x, y) \, dy,
\]
and
\[
\nabla \times \vec{D} \cdot \hat{n} dA = \left( \frac{\partial D_y}{\partial x} - \frac{\partial D_x}{\partial y} \right) \, dxdy.
\]
Using these two expressions the Stoke's theorem
\[
\oint_{\text{Curve bounding } A} \vec{D} \cdot d\vec{r} = \oiint_{\text{Surface } A} \nabla \times \vec{D} \cdot \hat{n} dA.
\]
(8.227)
can be expressed as
\[
\oint_{\text{Curve bounding } A} (D_x \, dx + D_y \, dy) = \oiint_{\text{Surface } A} \left( \frac{\partial D_y}{\partial x} - \frac{\partial D_x}{\partial y} \right) \, dxdy.
\]
(8.228)

Some Notes about Stokes' Theorem:
1. The curve \( C \) bounds the open area \( A \). The integral must therefore be the same for any area \( A \) bounded by the curve \( C \). This means that you can pick an easy area \( A \) if you wish to make life simpler!

2. The integral
\[
\int_A \nabla \times \bar{D} \cdot \hat{n} dA
\]  
(8.229)
is just the flux of the curl of the vector \( \bar{D} \) through the open area \( A \).

3. The integral
\[
\oint_C \bar{D} \cdot d\bar{r}
\]  
(8.230)
is the line integral of the vector \( \bar{D} \) around the closed curve \( C \). If \( \bar{D} \) is a force, then this integral is just the work done by that force around the curve \( C \). If the force is a conservative force, then the integral is zero.

Recall Green’s Theorem in the \( x-y \) Plane: Let’s consider two functions \( P(x, y) \) and \( Q(x, y) \). If these two functions and their first derivatives are continuous for all \( x \) and \( y \) in the region bounded by the curve \( C \) of area \( A \), then Green’s theorem states
\[
\iint_A \left[ \frac{\partial Q(x, y)}{\partial x} - \frac{\partial P(x, y)}{\partial y} \right] dx dy = \oint_C \left[ P(x, y) dx + Q(x, y) dy \right],
\]  
(8.231)
where the line integral is counterclockwise around the boundary of the area \( A \) (i.e. on the curve \( C \)). Note that for
\[
\bar{D} = D_x (x, y) \hat{x} + D_y (x, y) \hat{y} = P(x, y) \hat{x} + Q(x, y) \hat{y}
\]  
(8.232)
we find
\[
\iint_A \left[ \frac{\partial D_y}{\partial x} - \frac{\partial D_x}{\partial y} \right] dx dy = \oint_C [D_x dx + D_y dy].
\]  
(8.233)
This shows that Green’s theorem is the two-dimensional form of Stoke’s theorem.

**Example 8.8** Consider the vector field
\[
\bar{D}(x, y, z) = (3y, -xz, yz^2)
\]  
(8.234)
and the paraboloid
\[
2z = x^2 + y^2
\]  
(8.235)
bounded by the plane \( z = 2 \).

(a) Evaluate the flux integral directly
\[
\int_A \nabla \times \bar{D} \cdot \hat{n} dA
\]  
(8.236)
(b) Evaluate the flux integral indirectly by exploiting Stokes' theorem.

Solution:

(a) The curl of the vector \( \vec{D}(x, y, z) = (3y, -xz, yz^2) \) is given by

\[
\nabla \times \vec{D} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ D_x & D_y & D_z \end{vmatrix}
\]

(8.237)

becomes

\[
\nabla \times \vec{D} = \left( \frac{\partial (yz^2)}{\partial y} - \frac{\partial (-xz)}{\partial z} \right) \hat{x} + \left( \frac{\partial (3y)}{\partial z} - \frac{\partial (yz^2)}{\partial x} \right) \hat{y} + \left( \frac{\partial (-xz)}{\partial x} - \frac{\partial (3y)}{\partial y} \right) \hat{z} = (z^2 + x) \hat{x} - (3 + z) \hat{z}
\]

(8.238)

for \( z = 2 \), we find

\[
\nabla \times \vec{D} = (4 + x) \hat{x} - 5 \hat{z}
\]

(8.239)

Then the flux becomes

\[
\int_A \nabla \times \vec{D} \cdot \hat{n} dA = \int_A \left[ (4 + x) \hat{x} - 5 \hat{z} \right] \cdot \hat{n} \, dA
\]

\[
= \int_2^{-2} \int_{\sqrt{4-x^2}}^{\sqrt{4-x^2}} \left[ (4 + x) \hat{x} - 5 \hat{z} \right] \cdot dxdy
\]

(8.240)

which gives

\[
\int_A \nabla \times \vec{D} \cdot \hat{n} dA = -5 \int_{-2}^{2} \left( \int_{\sqrt{4-x^2}}^{\sqrt{4-x^2}} dy \right) \, dx = -10 \int_{-2}^{2} \sqrt{4-x^2} \, dx
\]

(8.241)
CHAPTER 8. VECTOR CALCULUS

Using the integral

\[
\ln[9]:= \int_{-2}^{2} \sqrt{4 - x^2} \, dx
\]

\[
\text{Out[9]} = 2 \pi
\]

we find

\[
\int_{A} \nabla \times \vec{D} \cdot \hat{n} dA = -20\pi. \quad (8.242)
\]

(b) Applying Stokes’ theorem

\[
\int_{\text{Surface } A} \nabla \times \vec{D} \cdot \hat{n} dA = \oint_{\text{Curve bounding } A} \vec{D} \cdot d\vec{r}. \quad (8.243)
\]

we have

\[
\vec{D}(x, y, z) \cdot d\vec{r} = (3y, -xz, yz^2) \cdot (dx, dy) = 3ydx - xzdy \quad (8.244)
\]

Using the circular curve of radius \( R = 2 \) on the plane \( z = 2 \) bounding the circular surface we considered in part a, we may write

\[
\int_{C} \vec{D} \cdot d\vec{r} = \int_{-2}^{2} \left[ 3 \left( -\sqrt{4 - x^2} \right) dx - 2x \left( -\sqrt{4 - x^2} \right) \right] \\
+ \int_{-2}^{2} \left[ 3 \left( \sqrt{4 - x^2} \right) dx - 2x \left( \sqrt{4 - x^2} \right) \right] \quad (8.245)
\]
\[
\oint_C \mathbf{D} \cdot d\mathbf{r} = - \int_{-2}^{2} \left[ 3\sqrt{4 - x^2} + \frac{2x^2}{\sqrt{4 - x^2}} \right] dx
+ \int_{2}^{2} \left[ 3\sqrt{4 - x^2} + \frac{2x^2}{\sqrt{4 - x^2}} \right] dx
\] 
(8.246)

\[
\oint_C \mathbf{D} \cdot d\mathbf{r} = - \int_{-2}^{2} \left[ 3\sqrt{4 - x^2} + \frac{2x^2}{\sqrt{4 - x^2}} \right] dx
- \int_{-2}^{2} \left[ 3\sqrt{4 - x^2} + \frac{2x^2}{\sqrt{4 - x^2}} \right] dx
\] 
(8.247)

\[
\oint_C \mathbf{D} \cdot d\mathbf{r} = -2 \int_{-2}^{2} \left[ 3\sqrt{4 - x^2} + \frac{2x^2}{\sqrt{4 - x^2}} \right] dx
\] 
(8.248)

By evaluating the integral we find
\[
\oint_C \mathbf{D} \cdot d\mathbf{r} = -20\pi
\] 
(8.249)

8.4.3 The Divergence Theorem

For a vector field \( \mathbf{V}(x, y, z) \)
\[
\iiint_{\text{volume}} \left( \nabla \cdot \mathbf{V}(x, y, z) \right) dV = \iint_{\text{Surface enclosing } \tau} \mathbf{V}(x, y, z) \cdot d\mathbf{n} \] 
(8.250)

Example 8.9 Consider the vector field
\[
\mathbf{V}(x, y, z) = (xy, yz, zx)
\] 
(8.251)

(a) Compute the flux \( \phi_v \) through the closed rectangular surface of sides \( a, b, \) and \( c \) shown in the figure by direct application of the definition of flux.

(b) Compute \( \phi_v \) through the same surface indirectly by applying the Divergence theorem.

Solution:

(a) The flux \( \phi_v \) for the vector field \( \mathbf{V}(x, y, z) \) over a given surface area \( A \) is defined as
\[
\phi_v = \iint_A \mathbf{V}(x, y, z) \cdot d\mathbf{n}.
\]

The figure above have three pairs of surfaces normal to the three axis. The flux through the surfaces normal to the \( x \) axis
\[
\phi_{vx} = \iint \mathbf{V}(a, y, z) \cdot \hat{x} dy dz + \iint \mathbf{V}(0, y, z) \cdot (-\hat{x}) dy dz
\] 
(8.252)
\[ \Phi_{vx} = \int_{0}^{c} \int_{0}^{b} \vec{V}(a,y,z) \cdot \hat{x} dy dz - \int_{0}^{c} \int_{0}^{b} \vec{V}(0,y,z) \cdot \hat{x} dy dz \quad (8.253) \]

\[ \Phi_{vx} = \int_{0}^{c} \int_{0}^{b} V_x(a,y,z) dy dz - \int_{0}^{c} \int_{0}^{b} V_x(0,y,z) dy dz \quad (8.254) \]

Noting that \( V_x(0,y,z) = 0 \) and \( V_x(a,y,z) = ay \), we find

\[ \Phi_{vx} = \int_{0}^{c} \int_{0}^{b} ay dy dz = \frac{ab^2c}{2}. \quad (8.255) \]

Similarly for the surfaces normal to the \( y \) axis, we may write

\[ \Phi_{vy} = \int \int \vec{V}(x,b,z) \cdot \hat{y} dx dz + \int \int \vec{V}(x,0,z) \cdot (-\hat{y}) dx dz \quad (8.256) \]

\[ \Rightarrow \Phi_{vy} = \int_{0}^{c} \int_{0}^{a} V_y(x,b,z) dx dz - \int_{0}^{c} \int_{0}^{a} V_y(x,0,z) dx dz. \quad (8.257) \]

Noting that \( V_y(x,0,z) = 0 \) and \( V_y(x,b,z) = bz \), we find

\[ \Phi_{vy} = \int_{0}^{c} \int_{0}^{a} bz dx dz = \frac{abc^2}{2}. \quad (8.258) \]

and

\[ \Phi_{vz} = \int_{0}^{c} \int_{0}^{a} cx dx dy = \frac{a^2bc}{2}. \quad (8.259) \]

Therefore, the total flux will be

\[ \Phi_v = \frac{a^2bc}{2} + \frac{ab^2c}{2} + \frac{abc^2}{2}. \quad (8.260) \]
(b) According to Divergence theorem
\[
\iiint_{\text{volume}} \left( \nabla \cdot \vec{V}(x,y,z) \right) \, dx \, dy \, dz = \iint_{\text{Surface enclosing } \tau} \vec{V}(x,y,z) \cdot \hat{n} \, da
\]
(8.261)

Thus the flux can be expressed as
\[
\phi_v = \iiint_{\text{volume}} \left( \nabla \cdot \vec{V}(x,y,z) \right) \, dx \, dy \, dz
\]
\[
= \int_0^c \int_0^b \int_0^a \left( \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} \right) \, dx \, dy \, dz
\]
\[
= \int_0^c \int_0^b \int_0^a \left( \frac{\partial (xy)}{\partial x} + \frac{\partial (yz)}{\partial y} + \frac{\partial (zx)}{\partial z} \right) \, dx \, dy \, dz
\]
\[
= \int_0^c \int_0^b \int_0^a (y + z + x) \, dx \, dy \, dz
\]
(8.262)

which leads to
\[
\phi_v = \frac{a^2 bc}{2} + \frac{ab^2 c}{2} + \frac{abc^2}{2}.
\]
(8.263)

8.4.4 More Examples on Divergence and Stokes theorems

The divergence in cylindrical coordinates:
\[
\nabla \cdot \vec{V} = \frac{1}{r} \frac{\partial}{\partial r} (r V_r) + \frac{1}{r} \frac{\partial}{\partial \varphi} (V_\varphi) + \frac{\partial}{\partial z} (V_z)
\]
(8.264)

**Example 8.10** Verify the divergence theorem for the vector field
\[
\vec{V} = ar \hat{r} + b \hat{\varphi} + cz \hat{z}
\]
(8.265)

(where \(a, b, \) and \(c\) are constants) and the cylinder shown below.

**Solution:**

Noting that
\[
V_r = ar, V_\varphi = b, V_z = cz
\]
(8.266)

the divergence
\[
\nabla \cdot \vec{V} = \frac{1}{r} \frac{\partial}{\partial r} (r V_r) + \frac{1}{r} \frac{\partial}{\partial \varphi} (V_\varphi) + \frac{\partial}{\partial z} (V_z)
\]
(8.267)

becomes
\[
\nabla \cdot \vec{V} = \frac{1}{r} \frac{\partial}{\partial r} (r^2 a) + \frac{1}{r} \frac{\partial}{\partial \varphi} (b) + \frac{\partial}{\partial z} (cz) = 2a + c
\]
(8.268)
Now applying the Divergence theorem

\[
\iiint_{\text{volume } \tau} \left( \nabla \cdot \vec{V}(x, y, z) \right) \, dx \, dy \, dz = \iint_{\text{Surface enclosing } \tau} \vec{V}(x, y, z) \cdot \hat{n} \, da
\]

(8.269)

For the volume integral, we find

\[
\iiint_{\text{volume } \tau} (2a + c) \, dx \, dy \, dz = (2a + c) \pi R^2 H.
\]

(8.270)

For the surface integral over the closed cylindrical surface we can write

\[
\iint_{\text{Surface enclosing } \tau} \vec{V}(r, \varphi, z) \cdot \hat{n} \, da = \iint_{\text{bottom}} \vec{V}_b(r, \varphi, z) \cdot \hat{n}_b \, da \\
+ \iint_{\text{top}} \vec{V}_t(r, \varphi, z) \cdot \hat{n}_t \, da + \iint_{\text{side}} \vec{V}_s(r, \varphi, z) \cdot \hat{n}_s \, da.
\]

(8.271)

We note that

\[
\hat{n}_b = -\hat{z}, \hat{n}_t = \hat{z}, \hat{n}_s = \hat{r}
\]

(8.272)

and

\[
\vec{V}_b(r, \varphi, z) = \vec{V}(r, \varphi, z = 0) = a\hat{r} + b\hat{\varphi}
\]

(8.273)

\[
\vec{V}_t(r, \varphi, z) = \vec{V}(r, \varphi, z = H) = a\hat{r} + b\hat{\varphi} + cH\hat{z}
\]

(8.274)
\[ \mathbf{V}_s (r, \phi, z) = \mathbf{V} (r = R, \phi, z) = a \mathbf{\hat{r}} + b \mathbf{\hat{\phi}} + c z \mathbf{\hat{z}} \] (8.275)

so that

\[
\iint_{\text{bottom}} \mathbf{V}_b (r, \phi, z) \cdot \mathbf{n}_b \, da = \iint_{\text{bottom}} (a \mathbf{\hat{r}} + b \mathbf{\hat{\phi}}) \cdot (-\mathbf{\hat{z}}) \, da = 0 \tag{8.276}
\]

\[
\iint_{\text{top}} \mathbf{V}_t (r, \phi, z) \cdot \mathbf{n}_t \, da = \iint_{\text{top}} (a \mathbf{\hat{r}} + b \mathbf{\hat{\phi}} + c H \mathbf{\hat{z}}) \cdot \mathbf{\hat{z}} \, da = \iint_{\text{top}} cH \, da = c \pi R^2 H \tag{8.277}
\]

and

\[
\iint_{\text{side}} \mathbf{V}_s (r, \phi, z) \cdot \mathbf{n}_s \, da = \iint_{\text{side}} (a \mathbf{\hat{r}} + b \mathbf{\hat{\phi}} + c z \mathbf{\hat{z}}) \cdot \mathbf{\hat{r}} \, da = \int_{\text{side}} aR \, da = aR (2\pi RH) \Rightarrow \iint_{\text{side}} \mathbf{V}_s (r, \phi, z) \cdot \mathbf{n}_s \, da = 2a \pi R^2 H. \tag{8.278}
\]

Therefore the integral over the closed cylindrical surface

\[
\iint_{\text{Surface \, enclosing \, } \tau} \mathbf{V} (r, \phi, z) \cdot \mathbf{n} \, da = \iint_{\text{bottom}} \mathbf{V}_b (r, \phi, z) \cdot \mathbf{n}_b \, da
\]
\[+ \iint_{\text{top}} \mathbf{V}_t (r, \phi, z) \cdot \mathbf{n}_t \, da + \iint_{\text{side}} \mathbf{V}_s (r, \phi, z) \cdot \mathbf{n}_s \, da. \tag{8.279}
\]

becomes

\[
\iint_{\text{Surface \, enclosing \, } \tau} \mathbf{V} (r, \phi, z) \cdot \mathbf{n} \, da = c \pi R^2 H + 2a \pi R^2 H. = (2a + c) \pi R^2 H
\]

\[
= \iiint_{\text{volume \, } \tau} \left( \nabla \cdot \mathbf{V} (x, y, z) \right) \, dx \, dy \, dz \tag{8.280}
\]

The Divergence in spherical Polar Coordinates:

\[
\nabla \cdot \mathbf{V} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 V_r) + \frac{1}{r \sin(\theta)} \frac{\partial}{\partial \theta} (\sin(\theta) V_\theta) + \frac{1}{r \sin(\theta)} \frac{\partial}{\partial \phi} (V_\phi) \tag{8.281}
\]

**Example 8.11** Verify the divergence theorem for the vector field

\[ \mathbf{V} = \mathbf{\hat{r}} \] (8.282)

and the sphere shown below.

**Solution:** For the flux, we find

\[
\iint_{\text{Surface}} \mathbf{V} (r, \theta, \phi) \cdot \mathbf{n} \, da = \int_0^{2\pi} \int_0^\pi R \cdot \mathbf{\hat{r}} R^2 \sin(\theta) \, d\theta \, d\phi
\]
\[
= \int_0^{2\pi} \int_0^\pi 4\pi R^3 \, d\theta \, d\phi
\]

\[= 4\pi R^3 \]
The divergence of the vector \( \vec{V} = r \hat{r} \)

\[
\nabla \cdot \vec{V} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 V_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta V_\theta) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \varphi} (V_\varphi)
\]

(8.283)

becomes

\[
\nabla \cdot \vec{V} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 V_r) = \frac{1}{r^3} \frac{\partial}{\partial r} (r^3) = 3
\]

(8.284)

Then the integral becomes

\[
\iiint_{\text{volume}} \left( \nabla \cdot \vec{V} (r, \theta, \varphi) \right) d\tau = \int_0^{2\pi} \int_0^\pi \int_0^3 r^2 \sin \theta \, dr \, d\theta \, d\varphi = 4\pi R^3.
\]

(8.285)

Example 8.12 Verify Stokes’ theorem for the magnetic field vector

\[
\vec{B} = 4\hat{r} + 3\hat{\theta} - 2\hat{\varphi}
\]

(8.286)

and the curve \( C \) shown in the diagram below.

Solution:

We are given the magnetic field in spherical coordinates with components

\[
B_r = 4, B_\theta = 3, B_\varphi = -2
\]

(8.287)
so that the curl of the magnetic field

\[
\nabla \times \vec{B} = \frac{1}{r \sin(\theta)} \left( \frac{\partial}{\partial \theta} (\sin(\theta) B_\phi) - \frac{\partial}{\partial \phi} (B_\theta) \right) \hat{r} \\
+ \left( \frac{1}{r \sin(\theta)} \frac{\partial}{\partial \phi} (B_r) - \frac{1}{r} \frac{\partial}{\partial r} (r B_\phi) \right) \hat{\theta} \\
+ \frac{1}{r} \left( \frac{\partial}{\partial r} (r B_\theta) - \frac{\partial}{\partial \theta} (B_r) \right) \hat{\phi}
\]

(8.288)

becomes

\[
\nabla \times \vec{B} = \frac{1}{r \sin(\theta)} \frac{\partial}{\partial \theta} (-2 \sin(\theta)) \hat{r} - \frac{1}{r} \frac{\partial}{\partial r} (-2r) \hat{\theta} + \frac{1}{r} \frac{\partial}{\partial r} (3r) \hat{\phi} \\
= -\frac{2 \cos(\theta)}{r \sin(\theta)} \hat{r} + \frac{2}{r} \hat{\theta} + \frac{3}{r} \hat{\phi} \Rightarrow \nabla \times \vec{B} = -\frac{2}{r} \cot(\theta) \hat{r} + \frac{2}{r} \hat{\theta} + \frac{3}{r} \hat{\phi} 
\]

(8.289)

Noting that the infinitesimal area on the x-y plane in spherical coordinates can be expressed as

\[
\hat{n} da = \hat{z} r \sin(\theta) dr d\phi = \left[ \cos(\theta) \hat{r} - \sin(\theta) \hat{\theta} \right] r \sin(\theta) dr d\phi
\]

(8.290)

then

\[
\nabla \times \vec{B} \cdot \hat{n} da = \left( -\frac{2}{r} \cot(\theta) \hat{r} + \frac{2}{r} \hat{\theta} + \frac{3}{r} \hat{\phi} \right) \left[ \cos(\theta) \hat{r} - \sin(\theta) \hat{\theta} \right] r \sin(\theta) dr d\phi
\]

(8.291)

\[
\Rightarrow \nabla \times \vec{B} \cdot \hat{n} da = \left( -\frac{2}{r} \cot(\theta) \cos(\theta) - \frac{2}{r} \sin(\theta) \right) r \sin(\theta) dr d\phi.
\]

(8.292)

Noting that on the x-y plane \( \theta = \pi/2 \), we have

\[
\nabla \times \vec{B} \cdot \hat{n} da = -2 r dr d\phi.
\]

(8.293)

so that the integral

\[
\int_A \nabla \times \vec{B} \cdot \hat{n} dA
\]

(8.294)
for the surface bounded by the curve shown in the figure can be written as
\[ \int_A \nabla \times \vec{B} \cdot \hat{n} dA = - \int_0^R \int_0^{\pi/2} 2r dr d\varphi = -R\pi. \] (8.295)

We recall the position vector in spherical coordinates is given by
\[ \vec{r} = r \hat{r} \Rightarrow d\vec{r} = dr \hat{r} + r d\hat{r}. \] (8.296)
where
\[ \hat{r} = \sin(\theta) \cos(\varphi) \hat{x} + \sin(\theta) \sin(\varphi) \hat{y} + \cos(\theta) \hat{z}. \] (8.297)
Since the closed curve is on the x-y plane (\( \theta = \pi/2 \)), we have
\[ \hat{r} = \cos(\varphi) \hat{x} + \sin(\varphi) \hat{y} \Rightarrow d\vec{r} = [-\sin(\varphi) \hat{x} + \cos(\varphi) \hat{y}] d\varphi = \hat{\varphi} d\varphi \] (8.298)
so that
\[ \vec{r} = r \hat{r} \Rightarrow d\vec{r} = dr \hat{r} + r d\hat{\varphi}. \] (8.299)
Then
\[ \vec{B} \cdot d\vec{r} = B_r dr + B_\varphi r d\varphi = 4dr - 2r d\varphi \] (8.300)
which leads to
\[ \oint_C \vec{B} \cdot d\vec{r} = \int_{l_1} (4dr - 2r d\varphi) + \int_{l_2} (4dr - 2r d\varphi) + \int_{l_3} (4dr - 2r d\varphi) \] (8.301)
where \( l_1 \) and \( l_3 \) are the straight lines (on this lines \( d\varphi = 0 \)) and \( l_2 \) is the curved line (on this line \( r = R \) and \( dr = 0 \)). Therefore
\[ \oint_C \vec{B} \cdot d\vec{r} = \int_0^R 4dr - \int_0^{\pi/2} 2R d\varphi + \int_0^0 4dr = -R\pi. \] (8.302)
The two results for the magnetic field vector
\[ \vec{B} = 4\hat{r} + 3\hat{\theta} - 2\hat{\varphi} \] (8.303)
verify that Stokes’ theorem
\[ \oint_C \vec{B} \cdot d\vec{r} = \int_A \nabla \times \vec{B} \cdot \hat{n} dA. \] (8.304)
is valid.

8.4.5 Applications of the Divergence and Stoke’s theorems

Gauss’ Law
\[ \nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \] (8.305)
**Gravity**

\[ \nabla \cdot \vec{g} = -4\pi G \rho \]  
(8.306)

**Ampere’s Law:** Ampere’s law states that

\[ \oint_C \vec{B} \cdot d\vec{r} = \mu_0 I_{\text{en}} \]  
(8.307)

where \( \vec{B} \) is the magnetic field due to the current \( I_{\text{en}} \) passing through the area enclosed by the curve \( C \) and \( \mu_0 \) is the magnetic permeability of free space. The differential form of Ampere’s law can be obtained if we apply Stokes’ theorem. Using Stokes’ theorem we can write

\[ \oint_C \vec{B} \cdot d\vec{r} = \int_A \nabla \times \vec{B} \cdot \hat{n}dA = \mu_0 I_{\text{en}} \]  
(8.308)

If the current passing through the area enclosed by the curve \( C \) is described by the current density \( \vec{J} \), the total current \( I_{\text{en}} \) can be expressed as

\[ I_{\text{en}} = \int_A \vec{J} \cdot \hat{n}dA \]  
(8.309)

so that

\[ \oint_C \vec{B} \cdot d\vec{r} = \int_A \nabla \times \vec{B} \cdot \hat{n}dA = \mu_0 I_{\text{en}} = \mu_0 \int_A \vec{J} \cdot \hat{n}dA \]

\[ \Rightarrow \int_A (\nabla \times \vec{B}) \cdot \hat{n}dA = \mu_0 \int_A \vec{J} \cdot \hat{n}dA \Rightarrow \nabla \times \vec{B} = \mu_0 \vec{J}. \]  
(8.310)

This is the differential form of Ampere’s law.

**Example 8.13** Who is wrong (Amperes or Stokes)? [From PHYS 4330]: Consider a parallel plate capacitor connected to a battery as shown in the figure below.

If we consider the flat circular area for the Amperian loop shown in the figure the current enclosed is just \( I_{\text{enc}} \)

\[ I_{\text{enc}} = \int_{\text{flat surface}} \vec{J} \cdot d\vec{a} = I. \]  
(8.311)

which leads to

\[ \oint_C \vec{B} \cdot d\vec{r} = \int_A \nabla \times \vec{B} \cdot \hat{n}dA \Rightarrow \int_A \vec{J} \cdot \hat{n}dA = \mu_0 I_{\text{en}} = \mu_0 I \]  
(8.312)

On the other hand, if we consider the balloon shaped surface shown in the figure there is no current passing through this surface and therefore the current enclosed is zero

\[ I_{\text{enc}} = \int_{\text{balloon surface}} \vec{J} \cdot d\vec{a} = 0. \]  
(8.313)
which leads to a different result

\[
\oint_C \vec{B} \cdot d\vec{r} = \int_A \nabla \times \vec{B} \cdot \hat{n} dA \\
\int_A \vec{J} \cdot \hat{n} dA = \mu_0 I_{en} = 0
\]  

(8.314)

Who is wrong here? Neither Ampere nor Stoke! However, in the case considered above the current is not steady because of the build up of charge on the capacitor. This suggested that Ampere’s Law need to be fixed so that it can also be used when the current is not steady and it was fixed by Maxwell. You want to learn more about this? TAKE THEORETICAL II (PHYS 3160) in spring 2013 and then E & M I & II (PHYS 4310 & 4330)!!!
Part II

Mathematical Methods in Physics II
Chapter 9

Introduction to the Calculus of Variations

9.1 Geodesic and stationary points

*Geodesic*: The curve along a surface which marks the shortest distance between two neighboring points. Finding geodesics is one of the problems which can be solved using the calculus of variation.

*Stationary point*: A point with coordinates, \((x_0, f(x_0))\), on a curve defined by the function \(f(x)\) is said to be a stationary point when

\[
\frac{df(x)}{dx}
|_{x=x_0} = 0.
\]  

(9.1)

Figure 9.1: Stationary points.

**Example 9.1** A ball of mass \(m\) is kicked from the ground level with an initial speed, \(v_0\), at an angle \(\theta_o\) above the horizontal. Find the time, \(t\), at which the height of the ball, \(y(t)\), becomes stationary.
Solution: We recall that from kinematics of a projectile the motion of the ball along the y direction is determined by Newton’s second law

\[
m \frac{dv_y}{dt} = -mg \Rightarrow \frac{dv_y}{dt} = -g = v_y(t) = v_{0y} - gt \quad (9.2)
\]

Noting that

\[
v_y(t) = \frac{dy(t)}{dt} \quad (9.3)
\]

the value of the time after the ball is kicked that makes the height function of the ball, \( y(t) \), stationary is given by

\[
v_y(t) = \frac{dy(t)}{dt} = 0 \Rightarrow t = \frac{v_{0y}}{g} = \frac{v_0 \sin (\theta_o)}{g} \quad (9.4)
\]

Example 9.2 Geodesic: Consider two points in a x-y plane \( P_1 \) and \( P_2 \). Prove that the shortest distance between the two points is the distance measured along a straight line (i.e. show that the geodesic is given by an equation of a straight line, \( y(x) = mx + b \)).

\[ P_2 (x_2, y_2) \]

\[ P_1 (x_1, y_1) \]

Figure 9.2: Geodesic and none geodesic paths.

Solution: Let’s consider two points on the x-y plane. Let \( P_1 \) be \((x_1, y_1)\) and \( P_2 \) be \((x_2, y_2)\). Then the distance between these points is given by the integral

\[
L = \int_{(1)}^{(2)} ds, \quad (9.5)
\]

where

\[
ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + \left( \frac{dy}{dx} \right)^2} \, dx = \sqrt{1 + \left( \frac{dx}{dy} \right)^2} \, dy. \quad (9.6)
\]
We may rewrite this distance as
\[ L = \int_{(1)}^{(2)} \sqrt{1 + \left( \frac{dy}{dx} \right)^2} \, dx. \tag{9.7} \]

Out of the infinitely many functions that can be used to connect the two points, we want to determine the one that would give the minimum distance. Let these function be denoted by \( Y(x) \). From these infinite number of functions there is only one function that gives the minimum distance between the two points. If this function is \( y(x) \), then we may write \( Y(x) \) in terms of \( y(x) \) as
\[ Y(x, \epsilon) = y(x) + \epsilon \eta(x), \tag{9.8} \]
where \( \eta(x) \) is an arbitrary function which must satisfy the condition
\[ \eta(x_1) = \eta(x_2) = 0 \tag{9.9} \]
so that at the two end points \( (x = x_1 = x_2) \), we find
\[ Y(x, \epsilon) = y(x). \tag{9.10} \]
Here \( \epsilon \) is the constant of variation. It is this constant that determines by how much \( Y(x) \) differs from \( y(x) \). Now in terms of \( Y(x) \), we may write
\[ L(\epsilon) = \int_{(1)}^{(2)} \sqrt{1 + Y'^2} \, dx, \tag{9.11} \]
where
\[ Y' = \frac{dY(x, \epsilon)}{dx}. \tag{9.12} \]
We are interested in the path that gives the minimum distance between the two points (i.e. the geodesic). The necessary condition for the distance, \( L(\epsilon) \), to be minimum is that the length function, \( L(\epsilon) \), must have a stationary point at \( (\epsilon = 0, L(\epsilon = 0)) \). This requires
\[ \left. \frac{dL(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = 0, \tag{9.13} \]
which leads to
\[ \left. \frac{dL(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = \left[ \int_{(1)}^{(2)} \left( -\frac{1}{2} \right) \frac{1}{\sqrt{1 + Y'^2}} (2Y') \left( \frac{dY'(x, \epsilon)}{d\epsilon} \right) dx \right]_{\epsilon=0} = 0. \tag{9.14} \]
Using
\[ Y(x, \epsilon) = y(x) + \epsilon \eta(x) \tag{9.15} \]
we may write
\[
\frac{dY}{dx} = \frac{dy}{dx} + \epsilon \frac{d\eta}{dx} \Rightarrow Y' (x, \epsilon) = y' (x) + \epsilon \eta' (x) \tag{9.16}
\]
so that
\[
\frac{dY'}{d\epsilon} = \frac{d}{d\epsilon} [y' (x) + \epsilon \eta' (x)] = \eta' (x). \tag{9.17}
\]
There follows that
\[
Y' (x, \epsilon)|_{\epsilon=0} = y' (x) \tag{9.18}
\]
and
\[
\frac{dY'}{d\epsilon} \bigg|_{\epsilon=0} = \eta' (x). \tag{9.19}
\]
In view of these results, one finds for the stationary point
\[
\frac{dL}{d\epsilon} \bigg|_{\epsilon=0} = \int_{(1)}^{(2)} \left(- \frac{1}{2}\right) \frac{1}{\sqrt{1 + y'^2}} \left(2Y' \right) \frac{dY'}{d\epsilon} \bigg|_{\epsilon=0} dx = 0. \tag{9.20}
\]
Using integration by parts
\[
\int u dv = uv - \int v du \tag{9.21}
\]
for
\[
\eta' (x) = dv \Rightarrow v = \eta (x), \tag{9.22}
\]
and
\[
u = \frac{y'}{\sqrt{1 + y'^2}} \Rightarrow du = \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + y'^2}}\right) dx, \tag{9.23}
\]
we may write the integral as
\[
\int_{(1)}^{(2)} \frac{y'\eta' (x)}{\sqrt{1 + y'^2}} dx = \left. \frac{y'}{\sqrt{1 + y'^2}} \eta (x) \right|_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta (x) \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + y'^2}}\right) dx = 0. \tag{9.24}
\]
Due to the conditions
\[
\eta (x_1) = \eta (x_2) = 0 \tag{9.25}
\]
the first term in the above expression becomes zero. Thus one can write
\[
\frac{dL}{d\epsilon} \bigg|_{\epsilon=0} = \int_{(1)}^{(2)} \frac{y'\eta' (x)}{\sqrt{1 + y'^2}} dx = - \int_{x_1}^{x_1} \eta (x) \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + y'^2}}\right) dx = 0. \tag{9.26}
\]
Since \( \eta(x) \) is an arbitrary function, for the integral to be zero, we must have

\[
\frac{d}{dx} \left( \frac{y'}{\sqrt{1+y'^2}} \right) = 0 \Rightarrow \frac{y'}{\sqrt{1+y'^2}} = c,
\]

(9.27)

where \( c \) is a constant. Upon solving for \( y' \)

\[
y' = \left[ \frac{c^2}{1-c^2} \right]^{1/2} = m.
\]

(9.28)

Note that we have introduced another constant in terms of the constant \( c \). There follows that

\[
\frac{dy}{dx} = m \Rightarrow y(x) = mx + b,
\]

(9.29)

which is equation of a straight line.

### 9.2 The general problem

In the previous section we saw the application of the calculus of variation to show that the shortest path connecting two points on a plane (the geodesic) is a straight line

\[
\frac{dy}{dx} = m \Rightarrow y(x) = mx + b.
\]

(9.30)

Next we shall consider the application of the calculus of variation to the general problem. In an Euclidean space a surface is defined by the function, \( F(x, y, z) \), where it depends on the Cartesian coordinates \( x, y, \) and \( z \). Instead of the Euclidean space let’s consider a surface defined by the function, \( F(x, y(x), y'(x) = \frac{dy}{dx}) \).

This surface could, for example, be a surface in phase space (in Classical mechanics) if we replace

\[
x \rightarrow t, y(x) \rightarrow y(t), y'(x) \rightarrow y'(t) = \frac{dy}{dt} = v_y(t) = \frac{p_y}{m}
\]

describing the dynamics of a particle mass, \( m \), moving along the \( y \)-direction in terms of the parameters \( \text{time} = t, \text{position} = y(t), \text{velocity} = v_y(t) \hat{y} = \frac{p_y}{m} \). where \( p_y \hat{y} \) is the momentum. In classical mechanics, the dynamics of a particle is determined by an equation derived from Newton’s second law. As we shall see, this equation can be derived from a more general equation known as the Euler-Lagrange Equation

\[
\frac{\partial}{\partial x} \left( \frac{\partial F}{\partial y'} \right) - \frac{\partial F}{\partial y} = 0,
\]

(9.31)

where \( F = F(x, y(x), y'(x)) \) is the function that defines the surface constructed by the set of points with coordinates, \((x, y(x), y'(x))\). The Euler-Lagrange Equation is derived by applying the calculus of variation. In general, in the problem...
that we want to solve applying the calculus of variation, we know the co-
ordinates of two different points \((x_1, y(x_1), y'(x_1))\) and \((x_2, y(x_2), y'(x_2))\) on the
surface defined by \(F = F(x, y(x), y'(x))\). From the infinitely many trajectories
that can connect these two points, there is only one trajectory on this surface
that is the shortest (the Geodesic). Finding the Geodesic is the general problem
that can be solved applying the calculus of variation.

The surface is defined by the function \(F(x, y(x), y'(x))\). The distance be-
tween these two points determined by evaluating the integral

\[
I = \int_{x_1}^{x_2} F(x, y(x), y'(x)) \, dx.
\]

(9.32)

To determine the equation that the function \(F\) is governed by so that we find the
shortest length joining the two points, let the function for any path connecting
the two points be \(Y(x)\). From these infinite number of functions there is only
one function that gives the minimum distance between the two points. If this
function is \(y(x)\), then we may write \(Y(x)\) in terms of \(y(x)\) as

\[
Y(x, \epsilon) = y(x) + \epsilon \eta(x),
\]

(9.33)

where \(\eta(x)\) is an arbitrary function which must satisfy the condition

\[
\eta(x_1) = \eta(x_2) = 0
\]

(9.34)

so that at the two points \((x = x_1 = x_2)\), we find

\[
Y(x, \epsilon) = y(x).
\]

(9.35)

We also have

\[
\frac{dY(x, \epsilon)}{d\epsilon} = \eta(x) \Rightarrow \left. \frac{dY(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = \eta(x),
\]

(9.36)
9.2. THE GENERAL PROBLEM

and

\[
dY &= \frac{dy}{dx} + \epsilon \frac{d\eta}{dx} \quad \text{or} \quad Y'(x, \epsilon) = y'(x) + \epsilon \eta'(x) \Rightarrow Y'(x, \epsilon)|_{\epsilon=0} = y'(x), \quad (9.37)
\]

which gives

\[
\frac{dY'(\epsilon)}{d\epsilon} = \frac{d}{d\epsilon} [y'(x) + \epsilon \eta'(x)] = \eta'(x) \Rightarrow \frac{dY'(\epsilon)}{d\epsilon}|_{\epsilon=0} = \eta'(x). \quad (9.38)
\]

For the Geodesic the integral

\[
I(\epsilon) = \int_{x_1}^{x_2} F(x, Y(x, \epsilon), Y'(x, \epsilon)) \, dx,
\]

must be stationary, that means

\[
\left. \frac{dI(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = \int_{x_1}^{x_2} \frac{d}{d\epsilon} \left[ F(x, Y(x, \epsilon), Y'(x, \epsilon)) \right] \left|_{\epsilon=0} \, dx = 0. \quad (9.40)
\]

Noting that

\[
\left. \frac{d}{d\epsilon} \left[ F(x, Y(x, \epsilon), Y'(x, \epsilon)) \right] \right|_{\epsilon=0} = \left. \frac{\partial F}{\partial Y} \frac{dY(\epsilon)}{d\epsilon} + \frac{\partial F}{\partial Y'} \frac{dY'(\epsilon)}{d\epsilon} \right|_{\epsilon=0}
\]

and substituting

\[
Y(x, \epsilon)|_{\epsilon=0} = y(x), \quad \left. \frac{dY(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = \eta(x), \quad Y'(x, \epsilon)|_{\epsilon=0} = y'(x), \quad \left. \frac{dY'(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = \eta'(x), \quad (9.42)
\]

we find

\[
\left. \frac{d}{d\epsilon} \left[ F(x, Y(x, \epsilon), Y'(x, \epsilon)) \right] \right|_{\epsilon=0} = \left[ \frac{\partial}{\partial y} F(x, y(x), y'(x)) \right] \eta(x) + \left[ \frac{\partial}{\partial y'} F(x, y(x), y'(x)) \right] \eta'(x), \quad (9.43)
\]

Then the integral for the Geodesic line becomes

\[
\left. \frac{dI(\epsilon)}{d\epsilon} \right|_{\epsilon=0} = \int_{x_1}^{x_2} \left[ \frac{\partial}{\partial y} F(x, y(x), y'(x)) \right] \eta(x) \, dx + \int_{x_1}^{x_2} \left[ \frac{\partial}{\partial y'} F(x, y(x), y'(x)) \right] \eta'(x) \, dx. \quad (9.44)
\]
Using integration by parts the second integral can be rewritten as
\[
\int_{x_1}^{x_2} \left[ \frac{\partial}{\partial y'} F(x, y(x), y'(x)) \right] \eta'(x) \, dx = \eta(x) \left. \frac{\partial F}{\partial y'} \right|_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta(x) \left. \frac{\partial F}{\partial y} \right|_{x_1} \frac{\partial y'}{\partial y'} \, dx,
\]
so that using
\[
\eta(x_1) = \eta(x_2) = 0,
\]
we find
\[
\int_{x_1}^{x_2} \left[ \frac{\partial}{\partial y'} F(x, y(x), y'(x)) \right] \eta'(x) \, dx = -\int_{x_1}^{x_2} \eta(x) \left. \frac{\partial F}{\partial y} \right|_{x_1} \eta(x) \, dx.
\]
(9.48)

Thus the stationary integral can be put in the form
\[
\frac{dI(\epsilon)}{d\epsilon} \bigg|_{\epsilon=0} = \int_{x_1}^{x_2} \left[ \frac{\partial}{\partial y} F(x, y(x), y'(x)) \right] \eta(x) \, dx + \int_{x_1}^{x_2} \eta(x) \left. \frac{\partial F}{\partial y} \right|_{x_1} \eta(x) \, dx = 0.
\]
(9.49)

or
\[
\frac{dI(\epsilon)}{d\epsilon} \bigg|_{\epsilon=0} = \int_{x_1}^{x_2} \left[ \frac{\partial}{\partial x} \left( \frac{\partial F}{\partial y'} \right) \right] \eta(x) \, dx = 0.
\]
(9.50)

There follow that
\[
\frac{\partial}{\partial x} \left( \frac{\partial F}{\partial y'} \right) - \frac{\partial F}{\partial y} = 0
\]
(9.51)

where \( F = F(x, y(x), y'(x)) \).

**Example 9.3** Do Example 9.2 using Euler-Lagrange Equation.

### 9.3 Applications: The Brachystochrone Problem

The Brachystochrone Problem: If two points A and B are given, at different heights but not lying one above the other (as shown in the figure below), it is required to find among all possible curves connecting them, that one along which a material point slides from A to B under the influence of gravity (neglecting friction) in the shortest possible time. This curve is called a Brachistochrone curve (Gr. βραχιστος, brachistos - the shortest, χρονος, chronos - time), or curve of fastest descent.[From Wikipedia, the free encyclopedia]

This problem occupied at the time of the leading mathematicians in the whole of Europe: Newton, Leibniz, Bernoulli, L’Hospital, and others. From then on, the calculus of variations developed as a special mathematical discipline.
Example 9.4 Using the Euler-Lagrange equation solve the brachystochrone problem, assuming the “material point” starts from rest.

Solution: We are given the two points $(x_1, y_1)$ and $(x_2, y_2)$; we chose axes through the point 1 with the $y$ axis positive downward as shown in Figure below. We want to find the curve joining the two points, down which a bead will slide (from rest) in the least time. That means we want to minimize time $t$. In other words we want to find the stationary value for the integral

$$I = \int_1^2 dt = \int_1^2 \frac{ds}{v}.$$  

The total energy of the bead is zero since it starts from rest assuming the zero energy level is the origin. If there is no friction, then we can write the energy at any point below the origin described by the coordinates $(x, y)$ as

$$\frac{1}{2}mv^2 - mgy = 0 \Rightarrow v = \sqrt{2gy}. \quad (9.52)$$

Then

$$I = \int_1^2 \frac{ds}{v} = \int_1^2 \frac{ds}{\sqrt{2gy}}. \quad (9.53)$$

Noting that

$$ds = \sqrt{dx^2 + dy^2}.$$
we have
\[
I = \int_1^2 \frac{\sqrt{dx^2 + dy^2}}{2g} = \int_1^2 \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{\sqrt{2g}} dy = \frac{1}{\sqrt{2g}} \int_{y_1}^{y_2} \frac{\sqrt{1 + x'^2}}{\sqrt{y}} dy,
\]
so that the stationary value of this integral is determined from the Euler-Lagrange equation
\[
\frac{\partial}{\partial y} \left( \frac{\partial F}{\partial x'} \right) - \frac{\partial F}{\partial x} = 0,
\]
where
\[
F (y, x(y), x'(y)) = \frac{1}{\sqrt{2g}} \sqrt{1 + x'^2},
\]
and
\[
x'(y) = \frac{dx(y)}{dy}.
\]
Noting that
\[
\frac{\partial F}{\partial x} = 0
\]
and
\[
\frac{\partial F}{\partial x'} = \frac{1}{\sqrt{2g} \sqrt{1 + x'^2} \sqrt{y}}
\]
so that
\[
\frac{\partial}{\partial y} \left( \frac{1}{\sqrt{2g} \sqrt{1 + x'^2} \sqrt{y}} \right) x' = \frac{x'}{\sqrt{1 + x'^2} \sqrt{y}} = \sqrt{c},
\]
where \(c\) is a constant. Solving for \(x'\), we find
\[
x' = \sqrt{c} \Rightarrow x'^2 = c (1 + x'^2) y
\]
\[
\Rightarrow x'^2 (1 - cy) = cy \Rightarrow \frac{dx}{dy} = \sqrt{\frac{cy}{1 - cy}} \Rightarrow x = \int_0^{-y} \sqrt{\frac{cy}{1 - cy}} dy.
\]
Introducing the transformation defined by
\[
cy = \sin^2 \left( \frac{\theta}{2} \right) = \frac{1}{2} (1 - \cos (\theta)) \Rightarrow dy = \frac{1}{c} \sin \left( \frac{\theta}{2} \right) \cos \left( \frac{\theta}{2} \right) d\theta
\]
we have
\[
x = \frac{1}{c} \int_0^{-y} \sqrt{\frac{cy}{1 - cy}} dy = \frac{1}{c} \int_0^{-\theta} \sqrt{\frac{\sin^2 \left( \frac{\theta}{2} \right)}{1 - \sin^2 \left( \frac{\theta}{2} \right)} \sin \left( \frac{\theta}{2} \right) \cos \left( \frac{\theta}{2} \right) d\theta
\]
\[
= \frac{1}{c} \int_0^{-\theta} \sin^2 \left( \frac{\theta}{2} \right) \sin \left( \theta \right) \cos \left( \frac{\theta}{2} \right) d\theta = \frac{1}{c} \int_0^{-\theta} \sin^2 \left( \frac{\theta}{2} \right) d\theta
\]
\[
\Rightarrow x = \frac{1}{c} \int_0^{-\theta} \frac{1}{2} (1 - \cos (\theta)) d\theta \Rightarrow x = \frac{1}{2c} (\theta - \sin (\theta)).
\]
Therefore, the trajectory of the bead that takes the smallest possible time is given by
\[ x = \frac{1}{2c} \left( \theta - \sin(\theta) \right), \quad y = \frac{1}{2c} \left( 1 - \cos(\theta) \right). \] (9.64)

Cycloid: Consider a circle of radius \( r \) rolling along the positive x-axis with a constant angular velocity starting from the origin. If you mark the point on the circle coinciding with the origin at the initial time (as shown in the Fig. 9.3)

![Figure 9.3: Cycloid motion.](image)

and follow the trajectory of this point its \( x \) and \( y \) coordinates of this point are given by
\[ x = r \left( \theta - \sin(\theta) \right), \quad y = r \left( 1 - \cos(\theta) \right), \] (9.65)

where \( \theta \) is the angle that the circle (the point) rotated. For example the figure below shows this trajectory for a point on a circle of unit radius \( (r = 1) \).

For a given \( \theta \), the circle’s centre lies at
\[ x = r\theta, \quad y = r. \] (9.66)

On the other hand if the circle’s center is
\[ x = r\theta, \quad y = -r. \] (9.67)

then the trajectory looks like the figure shown below

9.4 Applications: classical mechanics

Multivariable Functions: Suppose we are given a function, \( F \), that depends on \( y, z, dy/dx, dz/dx, \) and \( x \), and we want to find \( y = y(x) \) and \( z = z(x) \) which make
\[ I = \int_{x_1}^{x_2} F(x, y, z, y', z') \, dx \] (9.68)
stationary, then we must solve the Euler-Lagrange equations
\[
\frac{\partial}{\partial x} \left( \frac{\partial F}{\partial y'} \right) - \frac{\partial F}{\partial y} = 0, \quad (9.69)
\]
\[
\frac{\partial}{\partial x} \left( \frac{\partial F}{\partial z'} \right) - \frac{\partial F}{\partial z} = 0.
\]

9.5 Physical application of the Euler-Lagrange equation

1. The Hamiltonian \((H)\): The sum of the kinetic energy \((T)\) and potential energy \((V)\)
\[
H = T + V. \quad (9.70)
\]
2. The Lagrangian \((\mathcal{L})\): The kinetic energy minus the potential energy
\[
\mathcal{L} = T - V. \quad (9.71)
\]
3. The classical action \((S)\): the integral of the Lagrangian with respect to time over a given period of time
\[
S = \int_{t_1}^{t_2} \mathcal{L} dt. \quad (9.72)
\]
4. Hamilton’s Principle: The motion of a given system from time \(t_1\) to time \(t_2\) is such that the classical action
\[
S = \int_{t_1}^{t_2} \mathcal{L} dt \quad (9.73)
\]
9.5. PHYSICAL APPLICATION OF THE EULER-LAGRANGE EQUATION

has a stationary value for the correct path of the motion. The path actually followed by a system, as specified in terms of the generalized coordinates \( q_i \), is that path that makes the action integral stationary:

\[
I = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) \, dt = 0 \tag{9.74}
\]

for \( i = 1, 2, 3...n \). This means that the Lagrangian must satisfy the set of equations

\[
\frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \tag{9.75}
\]

for \( i = 1, 2, 3...n \).

**Example 9.5** Use Lagrange's equations to find the equation of motion for a particle traveling along the x-y plane under the influence of a potential energy function \( U(x) \).

**Solution:** The kinetic energy of a particle moving in the x-y plane can be expressed as

\[
T = \frac{1}{2} m \left( v_x^2 + v_y^2 \right) = \frac{1}{2} m \left( \dot{x}^2 + \dot{y}^2 \right). \tag{9.76}
\]

Then the Lagrangian

\[
L = T - U = \frac{1}{2} m \left( \dot{x}^2 + \dot{y}^2 \right) - U(x). \tag{9.77}
\]
Since $\mathcal{L} = \mathcal{L}(x, y, \dot{x}, \dot{y}, t)$ which is a function of two variables and we must have two Euler-Lagrange equations

$$\frac{\partial}{\partial t} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}} \right) - \frac{\partial \mathcal{L}}{\partial x} = 0, \quad \frac{\partial}{\partial t} \left( \frac{\partial \mathcal{L}}{\partial \dot{y}} \right) - \frac{\partial \mathcal{L}}{\partial y} = 0. \quad (9.78)$$

Therefore, using the Lagrangian we find

$$\frac{\partial}{\partial t} \left( \frac{\partial \mathcal{L}}{\partial \dot{x}} \right) - \frac{\partial \mathcal{L}}{\partial x} = 0 \Rightarrow \frac{\partial}{\partial t} (m \ddot{x}) + \frac{\partial U(x)}{\partial x} = 0$$

$$\Rightarrow m \ddot{x} = - \frac{\partial U(x)}{\partial x}$$

$$\frac{\partial}{\partial t} \left( \frac{\partial \mathcal{L}}{\partial \dot{y}} \right) - \frac{\partial \mathcal{L}}{\partial y} = 0 \Rightarrow m \ddot{y} = 0 \ (\text{Zero acceleration}) \quad (9.79)$$

**Example 9.6** The Atwood’s Machine: A string of length, $l$, passes over a frictionless pulley connecting two masses, $m_1$ and $m_2$. Find an expression for the acceleration of the masses in the system.

![Atwood's Machine](image-url)
Solution: Let’s define the origin of the $y$ axis on the surface of the ground also assume $m_1 > m_2$, and the length of the string is $l$. At a given time $t$ let the position of $m_1$ and $m_2$ be $y_1$ and $y_2$, as shown in Fig. 9.4. Then the kinetic energy of the system can be expressed as

$$T = \frac{1}{2}m_1\dot{y}_1^2 + \frac{1}{2}m_2\dot{y}_2^2$$  \hspace{1cm} (9.80)$$

and the gravitational potential energy

$$V = m_1gy_1 + m_2gy_2$$  \hspace{1cm} (9.81)$$

Then the Lagrangian

$$\mathcal{L} = T - V$$  \hspace{1cm} (9.82)$$

of the system can be written as

$$\mathcal{L} = \frac{1}{2}m_1\dot{y}_1^2 + \frac{1}{2}m_2\dot{y}_2^2 - m_1gy_1 - m_2gy_2.$$  \hspace{1cm} (9.83)$$

This equation appears to be a function of two variables. However, because of the constraint

$$y_1 + y_2 + l = C,$$  \hspace{1cm} (9.84)$$

where $C$ is a constant, we end up with a Lagrangian that depends only on one variable. If we replace

$$y_2 = C - y_1 - l \Rightarrow \dot{y}_2 = -\dot{y}_1$$  \hspace{1cm} (9.85)$$

we have

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)\dot{y}_1^2 - m_1gy_1 - m_2g(C - y_1 - l).$$  \hspace{1cm} (9.86)$$

which we may put in the form

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)\dot{y}_1^2 - (m_1 - m_2)gy_1 - C_1,$$  \hspace{1cm} (9.87)$$

where we replaced, $C_1 = m_2g(C - l)$. Now using

$$\frac{\partial}{\partial t} \left( \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = 0$$  \hspace{1cm} (9.88)$$

for $q_i = y_1$

$$\frac{\partial \mathcal{L}}{\partial y_1} = -(m_1 - m_2)g, \frac{\partial \mathcal{L}}{\partial \dot{y}_1} = (m_1 + m_2)\dot{y}_1$$  \hspace{1cm} (9.89)$$

we find

$$\frac{\partial}{\partial t} [(m_1 + m_2)\dot{y}_1] = -(m_1 - m_2)g \Rightarrow \ddot{a}_1 = \dot{y}_1 = \frac{m_1 - m_2}{m_1 + m_2}g$$  \hspace{1cm} (9.90)$$
Recalling that
\[ \dot{y}_2 = -\dot{y}_1 \] (9.91)
the acceleration of the second mass becomes
\[ a_1 = \frac{m_1 - m_2}{m_1 + m_2} g. \] (9.92)

The minus sign indicates the first mass is accelerating in the negative y-direction.

**Example 9.7** Central Forces: Describe the properties of the motion of a mass \( m \) moving under the influence of a central force (that is, a force acting only along the radial direction) given by
\[ \vec{F} = f(r) \hat{r} \] (9.93)
for some function \( f(r) \). Assume that the motion is confined to a plane.

**Solution:** The kinetic energy
\[ T = \frac{1}{2} m v^2. \] (9.94)
Using polar coordinates the magnitude of the velocity can be expressed as
\[ v = \dot{r}^2 + r^2 \dot{\theta}^2 \] (9.95)
and the kinetic energy becomes
\[ T = \frac{1}{2} m \left( \dot{r}^2 + r^2 \dot{\theta}^2 \right). \] (9.96)
The potential energy is related to the central force by
\[ \vec{F} = -\nabla \cdot U(r) \] (9.97)
where \( U(r) \) is the potential energy. Since the force is a central force it is directed along the radial direction and it depends on \( r \) only. Therefore the potential energy can be expressed as
\[ U(r) = -\int f(r) \, dr. \] (9.98)
Then the Lagrangian can be expressed as
\[ L \left( t, r, \dot{r}, \theta, \dot{\theta} \right) = T - U = \frac{1}{2} m \left( \dot{r}^2 + r^2 \dot{\theta}^2 \right) + \int f(r) \, dr \] (9.99)
Then using the Euler-Lagrange’s equation
\[ \frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \] (9.100)
we have
\[ \frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = 0, \frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = 0 \] (9.101)

so that using
\[\frac{\partial L}{\partial \theta} = 0, \frac{\partial L}{\partial \dot{\theta}} = m r^2 \ddot{\theta}, \frac{\partial L}{\partial r} = m r \ddot{\theta}^2 + f(r), \frac{\partial L}{\partial \dot{r}} = m \ddot{r} \] (9.102)

we find
\[\frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{\theta}} \right) = 0 \Rightarrow \frac{\partial L}{\partial \dot{\theta}} = \text{const} \Rightarrow m r^2 \ddot{\theta} = \text{cont} \Rightarrow I \omega = \text{cons.} \]

(Conservation of Ang. Mom.)
\[\frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = 0 \Rightarrow m \ddot{r} = m r \ddot{\theta}^2 + f(r) \] (9.103)
Chapter 10

Introduction to the
Eigenvalue Problem

10.1 Matrix Review

Matrix Arithmetic and Manipulation: Consider the following matrices:

\[
A = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 2 & 4 \\ 1 & -1 \\ 3 & -1 \end{pmatrix},
\]

\[
C = \begin{pmatrix} 2 & 1 & 3 \\ 4 & -1 & -2 \\ -1 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} -2 & 0 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & 0 \end{pmatrix} \tag{10.1}
\]

Multiplication by a Scalar: Any matrix can be multiplied by a scalar:

\[
2A = \begin{pmatrix} 2 \times 2 & 3 \times 2 & 1 \times 2 \\ 2 \times 2 & 1 \times 2 & 0 \times 2 \\ 4 \times 2 & 2 \times 2 & -8 \end{pmatrix},
2A = \begin{pmatrix} 4 & 6 & 2 \\ 4 & 2 & 0 \end{pmatrix} \tag{10.2}
\]

Addition and subtraction: Two matrices can be added or subtracted if and only if they have the same dimensions. From matrices \(A, B, C, \) and \(D\) we can add/subtract only matrices \(C\) and \(D\)

\[
C + D = \begin{pmatrix} 2 & 1 & 3 \\ 4 & -1 & -2 \\ -1 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -2 & 0 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & 0 \end{pmatrix}
= \begin{pmatrix} 2 - 2 & 1 + 0 & 3 + 1 \\ 4 + 1 & -1 - 1 & -2 + 2 \\ -1 + 3 & 0 + 1 & 1 + 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 4 \\ 5 & -2 & 0 \\ 2 & 1 & 1 \end{pmatrix}
\]
Matrix Multiplication: two matrices can be multiplied if and only if the number of columns of the first matrix is equal to the number of rows of the second matrix. If matrices have the same dimension, then they can be multiplied. From the above matrices we can make the multiplications:

\[
AB = \begin{pmatrix}
2 & 3 & 1 \\
2 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
2 & 4 \\
1 & -1 \\
3 & -1
\end{pmatrix} = \begin{pmatrix}
(ab)_{11} & (ab)_{12} \\
(ab)_{21} & (ab)_{22}
\end{pmatrix}
\] (10.3)

\[
CD = \begin{pmatrix}
2 & 1 & 3 \\
4 & -1 & -2 \\
-1 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
-2 & 0 & 1 \\
1 & -1 & 2 \\
3 & 1 & 0
\end{pmatrix} = \begin{pmatrix}
(cd)_{11} & (cd)_{12} & (cd)_{13} \\
(cd)_{21} & (cd)_{22} & (cd)_{23} \\
(cd)_{31} & (cd)_{32} & (cd)_{33}
\end{pmatrix}
\] (10.4)

but we can not make the matrix multiplications BC or BD.

The element in row \(i\) and column \(j\) of the product matrix \(AB\) is equal to row \(i\) of \(A\) times column \(j\) of \(B\). In index notation

\[
(ab)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}
\] (10.5)

where \((ab)_{ij}\) is the element of the product matrix \(AB\).

\* Commutativities: For any two multiplyable matrices \(C\) and \(D\),

\[
CD \neq DC
\] (10.6)

\* Commutator: For square matrices \(C\) and \(D\) the Commutator \([C, D]\) is defined as

\[
[C, D] = CD - DC
\] (10.7)

\* For any three matrices, \(F, G,\) and \(H\) that can be multiplied, we can write

The Associative Law:

\[
F(GH) = (FG)H
\] (10.8)

The Distributive Law:

\[
F(G + H) = FG + FH
\] (10.9)

\* The Identity Matrix; \(I\) (Boas: The Unit Matrix, \(U\))

\[
IA = AI = A
\] (10.10)

\* Transpose of a Matrix: The transpose of the matrix \(A\) is denoted by \(A^T\).

\[
A = \begin{pmatrix}
2 & 3 & 1 \\
2 & 1 & 0
\end{pmatrix}, \quad A^T = \begin{pmatrix}
2 & 2 \\
3 & 1 \\
1 & 0
\end{pmatrix}
\]

\[
B = \begin{pmatrix}
2 & 4 \\
1 & -1 \\
3 & -1
\end{pmatrix}, \quad B^T = \begin{pmatrix}
2 & 1 & 3 \\
4 & 1 & -1 \\
-1 & -1 & -1
\end{pmatrix}
\] (10.11)
10.1. MATRIX REVIEW

- **Adjoint of a Matrix**: The adjoint of a square matrix, $A$, is given by

$$
\text{adj}(A) = [\text{cof}(A)]^T
$$

where $\text{cof}(A)$ is the cofactor of the matrix $A$. We recall that the minor of matrix $A$, $(M_{ij})$, is the determinant of the matrix formed from matrix $A$ by removing the $i^{th}$ row and $j^{th}$ column. For the cofactor matrix the elements are expressed as

$$
[\text{cof}(A)]_{ij} = (-1)^{i+j} M_{ij}.
$$

- **Inverse of a (Square) Matrix**: $A^{-1}$

$$
A^{-1}A = AA^{-1} = I
$$

We can determine the inverse of an invertible matrix ($\det A \neq 0$) using row reduction or the adjoint matrix.

**a. Row reduction** in this approach for the matrix, for example,

$$
A = \begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{bmatrix}
$$

we start from

$$
\begin{bmatrix}
a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\
a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\
a_{31} & a_{32} & a_{33} & 0 & 0 & 1
\end{bmatrix}
$$

and do elementary row operation until we end up with

$$
\begin{bmatrix}
1 & 0 & 0 & b_{11} & b_{12} & b_{13} \\
0 & 1 & 0 & b_{21} & b_{22} & b_{23} \\
0 & 0 & 1 & b_{31} & b_{32} & b_{33}
\end{bmatrix}
$$

so that the inverse of the Matrix $A$ is given by

$$
A^{-1} = \begin{bmatrix}
b_{11} & b_{12} & b_{13} \\
b_{21} & b_{22} & b_{23} \\
b_{31} & b_{32} & b_{33}
\end{bmatrix}.
$$

**b. Using the adjoint matrix**: Using the adjoint matrix the inverse can be expressed as

$$
A^{-1} = \frac{[\text{cof}(A)]^T}{\det |A|}
$$

**Example 10.1** Find the inverse of the matrix

$$
A = \begin{bmatrix}
-1 & 2 & 3 \\
2 & 0 & -4 \\
-1 & -1 & 1
\end{bmatrix}
$$

using
a. Row reduction approach

b. The adjoint matrix approach

Solution: a. In the row reduction approach we start from

\[
\begin{bmatrix}
-1 & 2 & 3 & 1 & 0 & 0 \\
2 & 0 & -4 & 0 & 1 & 0 \\
-1 & -1 & 1 & 0 & 0 & 1
\end{bmatrix}
\]  
(10.21)

and try to get

\[
\begin{bmatrix}
1 & 0 & 0 & a_{11} & a_{12} & a_{13} \\
0 & 1 & 0 & a_{21} & a_{22} & a_{23} \\
0 & 0 & 1 & a_{31} & a_{32} & a_{33}
\end{bmatrix}
\]  
(10.22)

so that we can get the inverse matrix

\[
A^{-1} = \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{pmatrix}
\]  
(10.23)

b. Find the cofactor \( C_{ij} \) of the element \( A_{ij} \) in row \( i \) and column \( j \) which is equal to \((-1)^{i+j}\) times the value of the determinant remaining when we cross off row \( i \) and column \( j \). After you obtained all elements of the cofactor matrix write the cofactor matrix \( C \), transpose and divide it with the determinant of the matrix \( A \). The resulting matrix is \( A^{-1} \).

\[
A^{-1} = \frac{1}{|A|} C^T.
\]  
(10.24)

Ans:

\[
A^{-1} = -\frac{1}{2} \begin{pmatrix}
4 & 5 & 8 \\
-2 & -2 & -2 \\
2 & 3 & 4
\end{pmatrix}
\]  
(10.25)

10.2 Orthogonal matrices and the rotational operator

Matrices that make an orthogonal transformation of vectors. In an orthogonal transformation of vectors the magnitude of the vectors remains the same. For an orthogonal matrix

\[
M^{-1} = M^T
\]  
(10.26)

The Rotation Operator: For a counter-clockwise rotation about the z-axis by an angle \( \theta \), we denote the rotation matrix by \( R \): \( R_z(\theta) = R \). Then,

\[
r' = Rr
\]  
(10.27)

\[
\Rightarrow \begin{pmatrix}
x' \\
y' \\
z'
\end{pmatrix} = \begin{pmatrix}
\cos \theta & \sin \theta & 0 \\
-\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
\]
10.3 Eigenvalues and Eigenvectors

In general, a linear transformation of a vector, \( \vec{r} \), to a vector, \( \vec{r}' \), using matrix can be expressed as

\[
\vec{r}' = M \vec{r}.
\]  
(10.28)

In Cartesian coordinates

\[
\begin{pmatrix}
x' \\
y' \\
z'
\end{pmatrix} =
\begin{pmatrix}
M_{11} & M_{12} & M_{13} \\
M_{21} & M_{22} & M_{23} \\
M_{31} & M_{32} & M_{33}
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}.
\]  
(10.29)

Under this transformation if the vector \( \vec{r}' \) is expressible as

\[
\vec{r}' = M \vec{r} = \lambda \vec{r},
\]  
(10.30)

where \( \lambda \) is a constant, then the vector \( \vec{r} \) is called the eigenvector (characteristic vector) and \( \lambda \) is the eigenvalue (characteristic value) of the Matrix \( M \),

\[
\begin{pmatrix}
M_{11} & M_{12} & M_{13} \\
M_{21} & M_{22} & M_{23} \\
M_{31} & M_{32} & M_{33}
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix} = \lambda
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}.
\]  
(10.31)

Using the identity matrix, \( I \),

\[
I = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix},
\]  
(10.32)
one can put Eq. (10.31)
\[
\begin{pmatrix}
M_{11} & M_{12} & M_{13} \\
M_{21} & M_{22} & M_{23} \\
M_{31} & M_{32} & M_{33}
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= \lambda
\begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
\]
which can be rewritten as
\[
\left[
\begin{pmatrix}
M_{11} & M_{12} & M_{13} \\
M_{21} & M_{22} & M_{23} \\
M_{31} & M_{32} & M_{33}
\end{pmatrix}
- \left[
\begin{pmatrix}
\lambda & 0 & 0 \\
0 & \lambda & 0 \\
0 & 0 & \lambda
\end{pmatrix}
\right]
\right]
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= 0
\]
\Rightarrow
\begin{pmatrix}
M_{11} - \lambda & M_{12} & M_{13} \\
M_{21} & M_{22} - \lambda & M_{23} \\
M_{31} & M_{32} & M_{33} - \lambda
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= 0.
\]

The eigenvalues are obtained from the condition
\[
\begin{vmatrix}
M_{11} - \lambda & M_{12} & M_{13} \\
M_{21} & M_{22} - \lambda & M_{23} \\
M_{31} & M_{32} & M_{33} - \lambda
\end{vmatrix}
= 0,
\]
which is known as the eigenvalue equation (characteristic equation). To find the eigenvectors we substitute the eigenvalues and solve the resulting equations.

From The VNR Concise Encyclopedia of Mathematics (Van Nostrand Reinhold Co., publishers, 1977):

**Eigenvalues:** Eigenvalue problems are important in many branches of physics. They make it possible to find coordinate systems in which the transformations in question take on their simplest forms. In mechanics for instance, the principal moments of a rigid body are found with the help of the eigenvalues of the symmetric matrix representing the inertia tensor... Eigenvalues are of central importance in quantum mechanics, in which the measured values of physical "observables" appear as the eigenvalues of certain operators. The term "transformation" is used predominantly in pure mathematical (geometrical) context, whereas "operator" is more customary in applications (physics, technology).

**Example 10.2** Find the eigenvalues and the corresponding eigenvectors of the matrix
\[
M = \begin{pmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}
\]
10.3. EIGENVALUES AND EIGENVECTORS

Solution: The eigenvalue equation

\[
\begin{vmatrix}
0 - \lambda & 1 & 0 \\
1 & 0 - \lambda & 0 \\
0 & 0 & 0 - \lambda \\
\end{vmatrix} = 0 \Rightarrow
\begin{vmatrix}
-\lambda & 0 & 1 \\
0 & -\lambda & 0 \\
0 & 0 & -\lambda \\
\end{vmatrix} = 0
\]

\[\Rightarrow -\lambda^3 + \lambda = 0 \Rightarrow \lambda_1 = 0, \lambda_2 = 1, \lambda_3 = -1. \quad (10.37)\]

The corresponding eigenvectors are determined from

\[
\begin{pmatrix}
M_{11} - \lambda_i & M_{12} & M_{13} \\
M_{21} & M_{22} - \lambda_i & M_{23} \\
M_{31} & M_{32} & M_{33} - \lambda_i \\
\end{pmatrix}
\begin{pmatrix}
x_i \\
y_i \\
z_i \\
\end{pmatrix} = 0. \quad (10.38)
\]

For \(\lambda_1 = 0\), we find

\[
\begin{pmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
x_1 \\
y_1 \\
z_1 \\
\end{pmatrix} = 0 \Rightarrow x_1 = 0, y_1 = 0. \quad (10.39)
\]

We use the bra-ket notation to represent eigenvectors. The eigenvector for an eigenvalue, \(\lambda\), is represented using a ket-vector, \(|\lambda\rangle\), which is expressed as a column matrix

\[
|\lambda\rangle = \begin{pmatrix}
x \\
y \\
z \\
\end{pmatrix}. \quad (10.40)
\]

The eigenvector for \(\lambda_1\) can then be written as a column matrix

\[
|\lambda_1\rangle = \begin{pmatrix}
0 \\
0 \\
z_1 \\
\end{pmatrix}. \quad (10.41)
\]

The corresponding bra-vector, \(\langle \lambda |\), generally, is the transpose conjugate of the ket-vector. But here since we will consider only real values, the bra-vector is written as

\[
\langle \lambda_1 | = \begin{pmatrix}
x & y & z \\
\end{pmatrix}. \quad (10.42)
\]

Thus we can write

\[
\langle \lambda_1 | = \begin{pmatrix}
0 & 0 & z_1 \\
\end{pmatrix}. \quad (10.43)
\]

Both bra and ket vectors must be normalized (be unit vectors). Thus for any eigenvectors, we must have

\[
\langle \lambda | \lambda \rangle = 1.
\]

Normalizing this vector

\[
\langle \lambda_1 | \lambda_1 \rangle = 1 \Rightarrow \begin{pmatrix}
0 & 0 & z_1 \\
0 & 0 & z_1 \\
\end{pmatrix} = 1 \Rightarrow z_1 = 1 \Rightarrow |\lambda_1\rangle = \begin{pmatrix}
0 \\
0 \\
1 \\
\end{pmatrix}. \quad (10.44)
\]
Similarly for $\lambda_2 = 1$

\[
\begin{pmatrix}
-1 & 1 & 0 \\
1 & -1 & 0 \\
0 & 0 & -1
\end{pmatrix}
\begin{pmatrix}
x_2 \\
y_2 \\
z_2
\end{pmatrix} = 0 \Rightarrow -x_2 + y_2 = 0, x_2 - y_2 = 0, z_2 = 0
\]
\[
\Rightarrow x_2 = y_2, z_2 = 0
\]

(10.45)

and the eigenvector becomes

\[
|\lambda_2\rangle = x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},
\]

(10.46)

\[
\langle \lambda_2 | \lambda_2 \rangle = 1 \Rightarrow x_2^2 \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1 \Rightarrow x_2 = \frac{1}{\sqrt{2}}
\]

(10.47)

\[
\Rightarrow |\lambda_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}
\]

(10.48)

and for $\lambda_3 = -1$

\[
\begin{pmatrix}
1 & 1 & 0 \\
1 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
x_3 \\
y_3 \\
z_3
\end{pmatrix} = 0 \Rightarrow x_3 + y_3 = 0, x_3 + y_3 = 0, z_3 = 0
\]
\[
\Rightarrow x_3 = -y_3, z_3 = 0
\]

(10.49)

\[
\Rightarrow |\lambda_3\rangle = y_3 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \Rightarrow |\lambda_3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}
\]

(10.50)

Using Mathematica:
10.3. EIGENVALUES AND EIGENVECTORS

**Eigenvalue equation for Hermitian Matrices:** A Hermitian Matrix is a matrix that is equal to its transposed and conjugated matrix \( M^T = M \).

\[
\begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

These are the eigenvectors of \( M \).

\[
\{\lambda_2, \lambda_3, \lambda_4\} = \text{Eigenvalues}[\text{Matrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}]
\]

This lists the eigenvalues and the eigenvectors for \( M \). The eigenvalues are listed first.

\[
\text{Eigensystem}[\text{Matrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}]
\]

The Similarity Transformation: The similarity transformation of the matrix \( M \) is given by

\[
D = C^{-1}MC.
\]

The Similarity Transformation: The similarity transformation of the matrix \( M \) is given by

\[
D = C^{-1}MC.
\]

where \( C \) is a matrix whose columns are the eigenvectors of the Eigenvalue equation for matrix \( M \). The matrix \( D \) is a diagonal matrix where the diagonal elements are the eigenvalues.
10.4 Physical applications

Normal modes of Vibration for couple harmonic oscillators

Example 10.4 Consider a system consisting of two equal masses \( m \) connected by three identical springs of spring constant \( k \).

![Coupled harmonic oscillators](image)

Figure 10.1: Coupled harmonic oscillators.

The masses can slide on a horizontal, frictionless surface. The springs are at their unstretched/uncompressed lengths when the masses are at their equilibrium positions. At \( t = 0 \), the masses are displaced from their equilibrium positions by the amounts \( x_{10} \) and \( x_{20} \) and released from rest, as shown in the figure above. Completely describe the resulting motion.

Step 1: The Equations of Motion:

Step 2: Similarity Transformation (find the eigenvalues and eigenvectors)

Step 3: Solving the De-coupled Transformed Equations of Motion

Step 4: The Propagator Matrix, \( U \)

Step 5: The Normal Modes of Vibration

Solution:

Step 1 The Equations of Motion: To find the equations of motion I will use Euler-Lagrange equations. You can use Newton’s second law if you want to. To use the Euler-Lagrange equation we need to find the kinetic energy and the potential energy. Suppose at a given instant of time the position of the first mass is \( x_1 \) and the second mass is \( x_2 \) as measured from their corresponding equilibrium positions. Then the corresponding kinetic energy can be expressed as

\[
K_1 = \frac{1}{2}m\dot{x}_1^2, \quad K_2 = \frac{1}{2}m\dot{x}_2^2
\]  

(10.56)
so that the total kinetic energy be

\[ K = \frac{1}{2} m (\ddot{x}_1^2 + \ddot{x}_2^2). \]  

(10.57)

The potential energy (elastic) is due to the displacement of the springs from their equilibrium position. If the first mass is displaced by \( x_1 \) from the equilibrium and the second mass by \( x_2 \) in the positive \( x \) direction, then the first spring will be stretched by \( x_1 \), the second spring will be compressed by \( x_1 \) and at the same time it will be stretched by \( x_2 \); as a result the net displacement will be \( x_2 - x_1 \). The third spring will be compressed by \( x_2 \). Therefore, the total potential energy will be

\[ K = \frac{1}{2} k x_1^2 + \frac{1}{2} k x_2^2 + \frac{1}{2} k (x_2 - x_1)^2. \]  

(10.58)

Then the Lagrangian

\[ L(t, x_1, x_2, \dot{x}_1, \dot{x}_2) = T - U \]

\[ \Rightarrow L(t, x_1, x_2, \dot{x}_1, \dot{x}_2) = \frac{1}{2} m (\ddot{x}_1^2 + \ddot{x}_2^2) - \frac{1}{2} k x_1^2 - \frac{1}{2} k x_2^2 - \frac{1}{2} k (x_2 - x_1)^2 \]  

(10.59)

which leads to

\[ \frac{\partial L}{\partial x_1} = -kx_1 + k (x_2 - x_1), \quad \frac{\partial L}{\partial x_2} = -kx_2 - k (x_2 - x_1), \]

\[ \frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{x}_1} \right) = m \ddot{x}_1, \quad \frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{x}_2} \right) = m \ddot{x}_2. \]  

(10.60)

Using the Euler-Lagrange's equation

\[ \frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{x}_i} \right) - \frac{\partial L}{\partial x_i} = 0 \text{ for } i = 1, 2 \]  

(10.61)

we find

\[ m \ddot{x}_1 - [-kx_1 + k (x_2 - x_1)] = 0 \Rightarrow m \ddot{x}_1 = -2kx_1 + kx_2 \]  

(10.62)

and

\[ m \ddot{x}_2 - [-kx_2 - k (x_2 - x_1)] = 0 \Rightarrow m \ddot{x}_2 = -2kx_2 + kx_1. \]  

(10.63)

If we introduce a constant

\[ \omega = \sqrt{\frac{k}{m}}, \]  

(10.64)

then the above two equations can be put in the form

\[ \ddot{x}_1 = -2\omega^2 x_1 + \omega^2 x_2 \]  

(10.65)
and
\[ \ddot{x}_2 = \omega^2 x_1 - 2\omega^2 x_2. \] (10.66)

These two equations describe the equations of motion for the two masses. As you can see the two equations are coupled second order differential equations.

**Step 2: Similarity transformation:** In a matrix form the two equations can be put in the form
\[
\begin{pmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{pmatrix} =
\begin{pmatrix}
-2\omega^2 & \omega^2 \\
\omega^2 & -2\omega^2
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix}
\] (10.67)
or
\[ \ddot{\bar{r}} = M \bar{r} \] (10.68)

where
\[ \bar{r} = x_1 \hat{e}_1 + x_2 \hat{e}_2, \] (10.69)
and
\[ M = \begin{pmatrix}
-2\omega^2 & \omega^2 \\
\omega^2 & -2\omega^2
\end{pmatrix}. \] (10.70)

Note that \( \hat{e}_1 \) and \( \hat{e}_2 \) are column matrices given by
\[ \hat{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \hat{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \] (10.71)

Suppose if one solves the eigen value equation for the matrix \( M \)
\[ M |\lambda\rangle = \lambda |\lambda\rangle, \] (10.72)
we know that the matrix, \( C \), constructed from the normalized eigen vectors \( |\lambda\rangle \), leads to a similarity transformation for the matrix \( M \). That means
\[
C^{-1}MC = D
\]
\[ \Rightarrow \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \] (10.73)

where
\[ |\lambda_1\rangle = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, |\lambda_2\rangle = \begin{pmatrix} a_{11} \\ a_{22} \end{pmatrix} \] (10.74)
are the normalized eigen vectors. Noting that
\[ C^{-1}C = CC^{-1} = I \]
one can write the matrix for the equation of motion as
\[
\ddot{\mathbf{r}} = \mathbf{M}\ddot{\mathbf{r}} \Rightarrow \ddot{\mathbf{r}} = \mathbf{M}\mathbf{C}\mathbf{C}^{-1}\ddot{\mathbf{r}}
\]  
(10.75)
and multiplying from the left by the matrix \(\mathbf{C}^{-1}\), one can put the equation of motion in the form
\[
\mathbf{C}^{-1}\ddot{\mathbf{r}} = (\mathbf{C}^{-1}\mathbf{M})\mathbf{C}^{-1}\ddot{\mathbf{r}}
\]  
(10.76)
and introducing a new vector defined by
\[
\mathbf{C}^{-1}\mathbf{r} = \mathbf{C}^{-1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Rightarrow \mathbf{C}^{-1}\ddot{\mathbf{r}} = \mathbf{C}^{-1}\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = \begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix},
\]  
(10.77)
and noting that
\[
\mathbf{C}^{-1}\mathbf{M}\mathbf{C} = \mathbf{D}
\]
is a similarity transformation, we find
\[
\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.
\]
It is important to note that \(\mathbf{C}\) is a matrix whose columns are the eigenvectors of the Eigenvalue equation for matrix \(\mathbf{M}\). The matrix \(\mathbf{D}\) is a diagonal matrix in which the diagonal elements are the eigenvalues. Suppose if we can find the eigenvectors, \(\tilde{\mathbf{R}}\), such that
\[
\mathbf{M}\tilde{\mathbf{R}} = \lambda\tilde{\mathbf{R}},
\]  
(10.78)
or using the bra-ket notation (also known as Dirac notation)
\[
\mathbf{M}\left|\lambda\right> = \lambda\left|\lambda\right>,
\]  
(10.79)
then for the eigenvalue, one can written
\[
\det\begin{vmatrix} -2\omega^2 - \lambda & \omega^2 \\ \omega^2 & -2\omega^2 - \lambda \end{vmatrix} = 0 \Rightarrow (2\omega^2 + \lambda)^2 - (\omega^2)^2 = 0
\]  
\[\Rightarrow (2\omega^2 + \lambda - \omega^2)(2\omega^2 + \lambda + \omega^2) = 0 \Rightarrow \lambda_1 = -\omega^2, \lambda_2 = -3\omega^2.
\]  
(10.80)
The corresponding eigenvectors are obtained from
\[
\begin{pmatrix} -2\omega^2 - \lambda_1 & \omega^2 \\ \omega^2 & -2\omega^2 - \lambda_1 \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = 0
\]  
(10.81)
and
\[
\begin{pmatrix} -2\omega^2 - \lambda_2 & \omega^2 \\ \omega^2 & -2\omega^2 - \lambda_2 \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = 0.
\]  
(10.82)
Solving these equations for, $\lambda_1 = -\omega^2$, we find
\[
\begin{pmatrix}
-\omega^2 & \omega^2 \\
\omega^2 & -\omega^2
\end{pmatrix}
\begin{pmatrix}
X_1 \\
X_2
\end{pmatrix}
= 0 \Rightarrow -X_1 + X_2 = 0 \Rightarrow X_2 = X_1
\]
(10.83)
for, $\lambda_2 = -3\omega^2$,
\[
\begin{pmatrix}
\omega^2 & \omega^2 \\
\omega^2 & -\omega^2
\end{pmatrix}
\begin{pmatrix}
X_1 \\
X_2
\end{pmatrix}
= 0 \Rightarrow X_1 + X_2 = 0 \Rightarrow X_2 = -X_1.
\]
(10.84)
Then the two eigenvectors becomes
\[
\vec{R}_1 = X_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{R}_2 = X_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\]
(10.85)
Upon normalizing these eigen vectors, we find
\[
\hat{R}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \hat{R}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\]
(10.86)
or using bra-ket notation
\[
|\lambda_1 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, |\lambda_2 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\]
(10.87)
You can use Mathematica to check your results. Here is an example how to find eigenvalues and unnormalized eigenvectors determined using Mathematica:

\begin{verbatim}
In[3]:= M1 = {{-2 \omega^2, \omega^2}, {\omega^2, -2 \omega^2}} // MatrixForm
Out[3]//MatrixForm=
\begin{pmatrix}
-2 \omega^2 & \omega^2 \\
\omega^2 & -2 \omega^2
\end{pmatrix}

In[5]:= Eigenvalues[{{-2 \omega^2, \omega^2}, {\omega^2, -2 \omega^2}}]
Out[5]= {-3 \omega^2, -\omega^2}

In[6]:= Eigenvectors[{{-2 \omega^2, \omega^2}, {\omega^2, -2 \omega^2}}]
Out[6]= {{-1, 1}, {1, 1}}

In[7]:= Eigensystem[{{-2 \omega^2, \omega^2}, {\omega^2, -2 \omega^2}}]
Out[7]= {{-3 \omega^2, -\omega^2}, {{-1, 1}, {1, 1}}}
\end{verbatim}
Step 3: Solving the Decoupled Transformed Equations of Motion: The matrix $C$ and its inverse $C^{-1}$ can be expressed as

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}. \quad (10.88)$$

N.B. You must use the method we studied to find the inverse of the matrix $C$. But, here I am going to use mathematica to find the inverse

$$\text{In[8]}: \text{M2} = \begin{bmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \end{bmatrix} // \text{MatrixForm}$$

$$\text{Out[9]/MatrixForm} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\text{In[10]}: \text{Inverse}\left[\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}\right] // \text{MatrixForm}$$

$$\text{Out[10]/MatrixForm} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

We note that

$$CC^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \Rightarrow CC^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Rightarrow C^{-1}C = I \quad (10.89)$$

Recalling that

$$\ddot{\mathbf{r}} = M\ddot{\mathbf{f}} \quad (10.90)$$

can be written as

$$\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

where

$$C^{-1}\ddot{\mathbf{r}} = C^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Rightarrow C^{-1}\ddot{\mathbf{r}} = C^{-1} \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = \begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix}, \quad (10.91)$$

we can express the equations of motion as

$$\begin{pmatrix} \ddot{y}_1 \\ \ddot{y}_2 \end{pmatrix} = \begin{pmatrix} -\omega^2 & 0 \\ 0 & -3\omega^2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Rightarrow \ddot{y}_1 = -\omega^2 y_1, \ddot{y}_2 = -3\omega^2 y_2 \quad (10.92)$$
so that the solutions can be expressed as
\[ y_1 = A \cos(\omega t) + B \sin(\omega t), \quad y_2 = C \cos(\sqrt{3}\omega t) + D \sin(\sqrt{3}\omega t). \] \tag{10.93}

Now substituting these back into
\[ CC^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = C \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \] \tag{10.94}
we find
\[ \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} A \cos(\omega t) + B \sin(\omega t) \\ C \cos(\sqrt{3}\omega t) + D \sin(\sqrt{3}\omega t) \end{pmatrix} \] \tag{10.95}

Initially \((t = 0)\) the first mass is displaced, \(x_1(0) = x_{10}\) and the second mass displaced \(x_2(0) = x_{20}\), and both masses released from rest, \(x_1(0) = x_2(0) = 0\). Using these initial conditions we find
\[ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} \Rightarrow A = \frac{1}{\sqrt{2}} (x_{10} + x_{20}), C = \frac{1}{\sqrt{2}} (x_{10} - x_{20}) \tag{10.96} \]
and
\[ \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -\omega A \sin(\omega t) + \omega B \cos(\omega t) \\ -\sqrt{3}\omega C \sin\left(\sqrt{3}\omega t\right) + \sqrt{3}\omega D \cos\left(\sqrt{3}\omega t\right) \end{pmatrix} \Rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \omega B \\ \sqrt{3}\omega D \end{pmatrix} \Rightarrow B = 0, D = 0 \tag{10.97} \]

Then using the results above, one can write
\[ \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} (x_{10} + x_{20}) \cos(\omega t) \\ \frac{1}{\sqrt{2}} (x_{10} - x_{20}) \cos\left(\sqrt{3}\omega t\right) \end{pmatrix} \] \tag{10.98}

or
\[
x_1(t) = \frac{1}{2} (x_{10} + x_{20}) \cos(\omega t) + \frac{1}{2} (x_{10} - x_{20}) \cos\left(\sqrt{3}\omega t\right) \\
x_2(t) = \frac{1}{2} (x_{10} + x_{20}) \cos(\omega t) - \frac{1}{2} (x_{10} - x_{20}) \cos\left(\sqrt{3}\omega t\right) \tag{10.99} \]

**Step 4: The Propagator Matrix, U:** If upon simplifying the above expressions, we find
\[
x_1(t) = \frac{1}{2} \left[ \cos(\omega t) + \cos\left(\sqrt{3}\omega t\right) \right] x_{10} + \frac{1}{2} \left[ \cos(\omega t) - \cos\left(\sqrt{3}\omega t\right) \right] x_{20} \\
x_2(t) = \frac{1}{2} \left[ \cos(\omega t) - \cos\left(\sqrt{3}\omega t\right) \right] x_{10} + \frac{1}{2} \left[ \cos(\omega t) + \cos\left(\sqrt{3}\omega t\right) \right] x_{20} \tag{10.100} \]
which can be put using matrices as
\[
\begin{pmatrix}
    x_1(t) \\
    x_2(t)
\end{pmatrix}
= \frac{1}{2} \begin{pmatrix}
    \cos(\omega t) + \cos(\sqrt{3}\omega t) & \cos(\omega t) - \cos(\sqrt{3}\omega t) \\
    \cos(\omega t) - \cos(\sqrt{3}\omega t) & \cos(\omega t) + \cos(\sqrt{3}\omega t)
\end{pmatrix}
\begin{pmatrix}
    x_{10} \\
    x_{20}
\end{pmatrix},
\]
(10.101)
or
\[
\begin{pmatrix}
    x_1(t) \\
    x_2(t)
\end{pmatrix}
= U(t) \begin{pmatrix}
    x_1(0) \\
    x_2(0)
\end{pmatrix} \Rightarrow \vec{r} = U(t)\vec{r}(0),
\]
(10.102)
where
\[
\vec{r}(0) = \begin{pmatrix}
    x_1(0) \\
    x_2(0)
\end{pmatrix} = \begin{pmatrix}
    x_{10} \\
    x_{20}
\end{pmatrix},
\]
(10.103)
and
\[
U(t) = \frac{1}{2} \begin{pmatrix}
    \cos(\omega t) + \cos(\sqrt{3}\omega t) & \cos(\omega t) - \cos(\sqrt{3}\omega t) \\
    \cos(\omega t) - \cos(\sqrt{3}\omega t) & \cos(\omega t) + \cos(\sqrt{3}\omega t)
\end{pmatrix}
\]
is called the propagator matrix.

**Step 5: The Normal Modes of Vibration:** Suppose the initial state of the two masses is described by the first eigenvector. That means
\[
\vec{r}(0) = \begin{pmatrix}
    x_{10} \\
    x_{20}
\end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix}
    1 \\
    1
\end{pmatrix}
\]
(10.104)
then
\[
\vec{r} = U(t)\vec{r}(0)
\]
(10.105)
gives
\[
\begin{pmatrix}
    x_1(t) \\
    x_2(t)
\end{pmatrix}
= \frac{1}{2\sqrt{2}} \begin{pmatrix}
    \cos(\omega t) + \cos(\sqrt{3}\omega t) & \cos(\omega t) - \cos(\sqrt{3}\omega t) \\
    \cos(\omega t) - \cos(\sqrt{3}\omega t) & \cos(\omega t) + \cos(\sqrt{3}\omega t)
\end{pmatrix}
\begin{pmatrix}
    1 \\
    1
\end{pmatrix}
\]
(10.106)
which leads to
\[
\begin{pmatrix}
    x_1(t) \\
    x_2(t)
\end{pmatrix}
= \frac{1}{\sqrt{2}} \begin{pmatrix}
    \cos(\omega t) \\
    \cos(\omega t)
\end{pmatrix} \Rightarrow x_1(t) = x_2(t).
\]
(10.107)
The two masses oscillate with a frequency, \(\omega\), in the same direction. On the other hand, if the initially state of the two masses is given by the second eigenvector
\[
\vec{r}(0) = \begin{pmatrix}
    x_{10} \\
    x_{20}
\end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix}
    1 \\
    -1
\end{pmatrix}.
\]
(10.108)
then we have
\[
\begin{pmatrix}
  x_1(t) \\
  x_2(t)
\end{pmatrix} = \frac{1}{2\sqrt{2}} \begin{bmatrix}
  \cos(\omega t) + \cos(\sqrt{3}\omega t) & \cos(\omega t) - \cos(\sqrt{3}\omega t) \\
  \cos(\omega t) - \cos(\sqrt{3}\omega t) & \cos(\omega t) + \cos(\sqrt{3}\omega t)
\end{bmatrix}
\begin{bmatrix}
  1 \\
  -1
\end{bmatrix}
\]
(10.109)

which gives
\[
\begin{pmatrix}
  x_1(t) \\
  x_2(t)
\end{pmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix}
  \cos(\sqrt{3}\omega t) \\
  -\cos(\sqrt{3}\omega t)
\end{bmatrix} \Rightarrow x_1(t) = -x_2(t).
\]
(10.110)

The two masses oscillate with a frequency $\sqrt{3}\omega$ out of phase by $\pi$.

*The two modes of vibrations we saw above are called Normal Modes of vibration.*
Chapter 11

Special functions

11.1 The factorial and gamma function

The Factorial, $n!$: We recall the factorial function, $n!$, for a positive integer or zero is defined as

$$n! = n \times (n-1) \times (n-2) \times (n-3)(n-4)...3 \times 2 \times 1 \times 0!, \quad (11.1)$$

where

$$0! = 1. \quad (11.2)$$

The integral form of the Factorial function: Consider the integral function given by

$$F(p) = \int_0^\infty e^{-ax} \, dx. \quad (11.3)$$

For any real number, $\alpha > 0$, the value of this integral is

$$F(p) = \int_0^\infty e^{-ax} \, dx = \left. -\frac{e^{-ax}}{a} \right|_0^\infty = \frac{1}{\alpha}. \quad (11.4)$$

Now let’s differentiate this integral with respect to $\alpha$ as many as we can, say $n$ times (i.e. $\frac{\partial^n}{\partial \alpha^n}$). For the first derivative, $n = 1$

$$\int_0^\infty e^{-ax} \, dx = \frac{1}{\alpha} \Rightarrow \int_0^\infty x^0 e^{-ax} \, dx = \frac{0!}{\alpha^1}$$

$$\Rightarrow \frac{\partial}{\partial \alpha} \left[ \int_0^\infty e^{-ax} \, dx \right] = \frac{1}{\alpha} \Rightarrow \int_0^\infty \frac{\partial}{\partial \alpha} (e^{-ax}) \, dx = \frac{\partial}{\partial \alpha} \left( \frac{1}{\alpha} \right)$$

$$\int_0^\infty -xe^{-ax} \, dx = -\frac{1}{\alpha^2} \quad (11.5)$$

This can be put in the form

$$\int_0^\infty x^1 e^{-ax} \, dx = \frac{1!}{\alpha^2}. \quad (11.6)$$
CHAPTER 11. SPECIAL FUNCTIONS

For the second derivative \( n = 2 \)

\[
\frac{\partial^2}{\partial \alpha^2} \left( \int_0^\infty e^{-\alpha x} \, dx \right) = \frac{\partial^2}{\partial \alpha^2} \left( \frac{1}{\alpha} \right) \Rightarrow \frac{\partial}{\partial \alpha} \left( \int_0^\infty -xe^{-\alpha x} \, dx \right) = \frac{\partial}{\partial \alpha} \left( -\frac{11!}{\alpha^2} \right)
\]

\[
\Rightarrow \int_0^\infty x^2e^{-\alpha x} \, dx = \frac{2 \times 1!}{\alpha^3} = \frac{2!}{\alpha^3}.
\] (11.7)

For the third derivative \( n = 3 \)

\[
\frac{\partial^3}{\partial \alpha^3} \left( \int_0^\infty e^{-\alpha x} \, dx \right) = \frac{\partial^3}{\partial \alpha^3} \left( \frac{1}{\alpha} \right) \Rightarrow \frac{\partial}{\partial \alpha} \left( \int_0^\infty x^2e^{-\alpha x} \, dx \right) = \frac{\partial}{\partial \alpha} \left( \frac{2!}{\alpha^3} \right)
\]

\[
\Rightarrow \int_0^\infty -x^3e^{-\alpha x} \, dx = -\frac{3 \times 2 \times 1!}{\alpha^3} \Rightarrow \int_0^\infty x^3e^{-\alpha x} \, dx = \frac{3!}{\alpha^3}.
\] (11.8)

Therefore it is not difficult to see for the \( n \)th derivative, we find

\[
\int_0^\infty x^n e^{-\alpha x} \, dx = \frac{n!}{\alpha^n}.
\] (11.9)

Now we set \( \alpha = 1 \), we find

\[
n! = \int_0^\infty x^n e^{-x} \, dx,
\] (11.10)

which is the integral form of the Factorial function which is valid for an integer, \( n \geq 0 \).

The Gamma function: The Gamma function is defined by the integral Function given by

\[
\Gamma(p) = \int_0^\infty x^{p-1}e^{-x} \, dx.
\] (11.11)

where \( p > 0 \) is any positive real number. For \( p = n + 1 \), with \( n \geq 0 \) (positive integer or zero), we find

\[
\Gamma(n + 1) = \int_0^\infty x^n e^{-x} \, dx,
\] (11.12)

which is the Factorial function. Therefore, the factorial function in terms of the Gamma function can be expressed as

\[
n! = \Gamma(n + 1) = \int_0^\infty x^n e^{-x} \, dx.
\] (11.13)

The Recursion Relation: If we replace \( p \) by \( p + 1 \) in the expression for the Gamma function

\[
\Gamma(p) = \int_0^\infty x^{p-1}e^{-x} \, dx.
\] (11.14)

we find

\[
\Gamma(p + 1) = \int_0^\infty x^p e^{-x} \, dx.
\] (11.15)
11.1. THE FACTORIAL AND GAMMA FUNCTION

If we denote
\[
  u = x^p, \quad dv = e^{-x}dx \Rightarrow du = px^{p-1}, \quad v = -e^{-x},
\]
then using integration by parts
\[
  \int u dv = uv - \int v du
\]
we find
\[
  \Gamma(p + 1) = -e^{-x}x^p|_0^\infty - \int_0^\infty px^{p-1}(-e^{-x})dx.
\]
In the above expression the first term is zero
\[
  \Gamma(p + 1) = p \int_0^\infty x^{p-1}e^{-x}dx.
\]
We know that
\[
  \Gamma(p) = \int_0^\infty x^{p-1}e^{-x}dx
\]
hence
\[
  \Gamma(p + 1) = p\Gamma(p).
\]

**Example 11.1** Show that
\[
  \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}
\]
and
\[
  \int_0^\infty e^{-u^2} du = \frac{\Gamma\left(\frac{1}{2}\right)}{2}.
\]

**Solution:** Using the definition of the Gamma function, we can write
\[
  \Gamma\left(\frac{1}{2}\right) = \int_0^\infty \frac{1}{\sqrt{x}}e^{-x}dx.
\]
Introducing a new variable defined by
\[
  u^2 = x \Rightarrow 2udu = dx
\]
we can write
\[
  \Gamma\left(\frac{1}{2}\right) = \int_0^\infty \frac{1}{u}e^{-u^2}2udu = 2 \int_0^\infty e^{-u^2}du.
\]
Squaring both sides, we have
\[
  \Gamma^2\left(\frac{1}{2}\right) = 4 \left( \int_0^\infty e^{-u^2} du \right) \left( \int_0^\infty e^{-v^2} dv \right) = 4 \int_0^\infty \int_0^\infty e^{-(u^2+v^2)}dudv
\]
so that introducing the polar coordinates defined by

\[ u = r \cos \theta, \quad v = r \sin (\theta) \quad \Rightarrow \quad du \, dv = r \, dr \, d\theta, \quad u^2 + v^2 = r^2 \]

(11.28)
as shown in the figure below.

The above integral can be put in the form

\[ \Gamma^2 \left( \frac{1}{2} \right) = 4 \int_0^\infty e^{-r^2} r \, dr \int_0^{\pi/2} d\theta. \]

(11.29)

Here the set the upper limit of integration for \( \theta \) be \( \pi/2 \) since both \( u \) and \( v \) are positive and we must integrate only in the first quadrant (the region shown in green). Therefore, integrating with respect to \( r \) and \( \theta \) leads to

\[ \Gamma^2 \left( \frac{1}{2} \right) = 4 - \left. \frac{e^{-r^2}}{2} \right|_0^\infty \pi = \pi \Rightarrow \Gamma \left( \frac{1}{2} \right) = \sqrt{\pi}. \]

(11.30)

Now substituting this result into

\[ \Gamma \left( \frac{1}{2} \right) = \int_0^\infty \frac{1}{\sqrt{x}} e^{-x} \, dx = 2 \int_0^\infty e^{-u^2} \, du, \]

(11.31)

we can see that

\[ \int_0^\infty e^{-u^2} \, du = \sqrt{\frac{\pi}{2}}. \]

(11.32)
11.2 The Beta Function

For \( p > 0 \) and \( q > 0 \), the beta function \( B(p, q) \) is defined by a definite integral

\[
B(p, q) = \int_0^1 x^{p-1} (1 - x)^{q-1} \, dx.
\] (11.33)

There are different forms of representations of the beta function. These includes the following

1. Replace \( x = y/a \) \( \Rightarrow \) \( dx = dy/a \)

\[
B(p, q) = \frac{1}{a} \int_0^a \left( \frac{y}{a} \right)^{p-1} \left( 1 - \frac{y}{a} \right)^{q-1} \, dy = \frac{1}{a^{p+q-1}} \int_0^a y^{p-1} (a - y)^{q-1} \, dy.
\] (11.34)

2. Replace \( x = \sin^2(\theta) \) \( \Rightarrow \) \( dx = 2 \sin(\theta) \cos(\theta) \, d\theta \)

which gives

\[
B(p, q) = \int_0^{\pi/2} \sin^{2p-2}(\theta) (1 - \sin^2(\theta))^{q-1} 2 \sin(\theta) \cos(\theta) \, d\theta
\]

\[
\Rightarrow \quad B(p, q) = 2 \int_0^{\pi/2} \sin^{2p-1}(\theta) \cos^{2q-1}(\theta) \, d\theta \quad (11.35)
\]

3. Replacing \( x = y/(1 + y) \), we have

\[
dx = \frac{dy}{1 + y} - \frac{y \, dy}{(1 + y)^2} = \frac{dy}{(1 + y)^2}
\]

and

\[
x = 0 \Rightarrow y = 0, \quad x = 1 \Rightarrow \frac{y}{1 + y} = 1 \Leftrightarrow \lim_{y \to \infty} \frac{y}{1 + y} = 1 \quad (11.36)
\]

so that

\[
B(p, q) = \int_0^\infty \left( \frac{y}{1 + y} \right)^{p-1} \left( 1 - \frac{y}{1 + y} \right)^{q-1} \, dy
\]

\[
\Rightarrow \quad B(p, q) = \int_0^\infty \frac{y^{p-1} \, dy}{(1 + y)^{p+q}} \quad (11.37)
\]

**Example 11.2** Prove that the Gamma and the Beta Functions are related by

\[
B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)}. \quad (11.38)
\]
Solution: For the Gamma functions
\[ \Gamma(q) = \int_{0}^{\infty} x^{q-1} e^{-x} dx, \quad \Gamma(p) = \int_{0}^{\infty} y^{p-1} e^{-y} dy \]  
(11.39)
introducing the transformation of variables defined by
\[ u^2 = x \Rightarrow 2udu = dx, \quad v^2 = y \Rightarrow 2vdv = dy \]  
(11.40)
we find
\[ \Gamma(q) = 2 \int_{0}^{\infty} u^{2q-1} e^{-u^2} du, \quad \Gamma(p) = 2 \int_{0}^{\infty} v^{2p-1} e^{-v^2} dv \]  
(11.41)
Multiplying the two functions, we have
\[ \Gamma(p)\Gamma(q) = 4 \int_{0}^{\infty} \int_{0}^{\infty} u^{2q-1} v^{2p-1} e^{-(u^2+v^2)} du dv, \]  
(11.42)
so that using the polar coordinates
\[ u = r \cos \theta, \quad v = r \sin \theta \Rightarrow dudv = rdrd\theta, \quad u^2 + v^2 = r^2 \]  
(11.43)
we find
\[ \Gamma(p)\Gamma(q) = 4 \int_{0}^{\infty} \int_{0}^{\pi/2} (r \cos \theta)^{2q-1} (r \sin \theta)^{2p-1} e^{-r^2} rdrd\theta \]  
(11.44)
This can be rewritten as
\[ \Gamma(p)\Gamma(q) = 4 \int_{0}^{\infty} r^{2p+2q-3} e^{-r^2} rdr \int_{0}^{\pi/2} (\sin \theta)^{2p-1} (\cos \theta)^{2q-1} d\theta \]  
(11.45)
The first integral
\[ 4 \int_{0}^{\infty} r^{2p+2q-3} e^{-r^2} rdr = 2 \int_{0}^{\infty} R^{2p+2q-1} e^{-R} dR = 2 \Gamma(p + q) \]  
(11.46)
and applying
\[ B(p, q) = 2 \int_{0}^{\pi/2} \sin^{2p-1} (\theta) \cos^{2q-1} (\theta) d\theta \]  
(11.47)
we note that the second integral can be expressed in terms of the Beta function as
\[ \int_{0}^{\pi/2} (\sin \theta)^{3q-1} (\cos \theta)^{2p-1} d\theta = \frac{B(q, p)}{2}. \]  
(11.48)
Therefore
\[ B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)}. \]  
(11.49)
This equation relates the beta and gamma functions.
11.3 Stirling’s Formula

We recall the Gamma Function

\[ \Gamma(p) = \int_0^\infty x^{p-1} e^{-x} \, dx, \quad (11.50) \]

when \( p \) is zero or positive integer, gives the factorial function

\[ p! = \Gamma(p+1) = \int_0^\infty x^p e^{-x} \, dx. \quad (11.51) \]

Next we want to find an approximate formula when \( p \) is very large. This approximate formula is known as Stirling’s formula and is given by

\[ p! = (p+1) \approx p^p e^{-p} \sqrt{2\pi p}, \quad (11.52) \]

or if we take the natural logarithm of both sides

\[ \ln(p!) \approx \ln \left( p^p e^{-p} \sqrt{2\pi p} \right) = \ln(p^p) + \ln(e^{-p}) + \ln \left( (2\pi p)^{\frac{1}{2}} \right) \]

\[ = p \ln(p) - p \ln(e) + \frac{1}{2} \ln(2\pi p) \Rightarrow \ln(p!) \approx p \ln(p) - p + \frac{1}{2} \ln(2\pi p) \]

If \( p \) is very large, the last term is very small as compared to the first two terms and the Stirling’s formula is given by

\[ \ln(p!) \approx p \ln(p) - p. \quad (11.53) \]

**Proof:** Introducing a new variable defined by

\[ x = p + y\sqrt{p} \Rightarrow dx = \sqrt{p} \, dy, \quad x = 0 \Rightarrow y = -\sqrt{p}, \quad x \to \infty \Rightarrow y \to \infty, \quad (11.54) \]

the \( \Gamma \) function

\[ p! = \Gamma(p+1) = \int_0^\infty x^p e^{-x} \, dx \quad (11.55) \]

can be put in the form

\[ p! = \int_{-\sqrt{p}}^{\infty} (p + y\sqrt{p})^p e^{-(p+y\sqrt{p})} \sqrt{p} \, dy. \quad (11.56) \]

Noting that

\[ (p + y\sqrt{p})^p = e^{\ln[(p+y\sqrt{p})^p]} = e^{p \ln(p+y\sqrt{p})}, \quad (11.57) \]

one can write

\[ p! = \int_{-\sqrt{p}}^{\infty} e^{p \ln(p+y\sqrt{p})} e^{-(p+y\sqrt{p})} \sqrt{p} \, dy = \sqrt{p} \int_{-\sqrt{p}}^{\infty} e^{p \ln(p+y\sqrt{p}) - p-y\sqrt{p}} \, dy. \quad (11.58) \]
We recall that the Taylor series expansion for \( f(y) \) about \( y = 0 \) is given by
\[
  f(y) = f(0) + \frac{1}{1!} \frac{df(y)}{dy} \bigg|_{y=0} y + \frac{1}{2!} \frac{d^2 f(y)}{dy^2} \bigg|_{y=0} y^2 + \ldots \tag{11.59}
\]
so that for \( f(y) = \ln(p + y\sqrt{p}) \), using
\[
  f(0) = \ln(p) \quad \frac{df(y)}{dy} \bigg|_{y=0} = \frac{\sqrt{p}}{p + y\sqrt{p}} \bigg|_{y=0} = \frac{\sqrt{p}}{p},
\]
\[
  \frac{d^2 f(y)}{dy^2} \bigg|_{y=0} = \frac{d}{dy} \left[ \frac{\sqrt{p}}{p + y\sqrt{p}} \right] \bigg|_{y=0} = -\frac{\sqrt{p}\sqrt{p}}{(p + y\sqrt{p})^2} \bigg|_{y=0} = \frac{1}{p}, \tag{11.60}
\]
we find an approximate expression
\[
  \ln(p + y\sqrt{p}) \approx \ln(p) + \frac{\sqrt{p}y}{p} - \frac{y^2}{2p}. \tag{11.61}
\]

Then the approximate expression for the factorial becomes
\[
  p! \approx \sqrt{p} \int_{-\sqrt{p}}^{\infty} e^{p\ln(p+y\sqrt{p})-p-y\sqrt{p}} dy
\]
\[
  = \sqrt{p} \int_{-\sqrt{p}}^{\infty} e^{p\left\{\ln(p)+\frac{\sqrt{p}y}{p} - \frac{y^2}{2p}\right\}-y\sqrt{p}} dy = \sqrt{p} \int_{-\sqrt{p}}^{\infty} e^{p\left(\ln(p)-p-\frac{y^2}{2p}\right)} dy
\]
\[
  = \sqrt{pe^{p\ln(p)-p}} \int_{-\sqrt{p}}^{\infty} e^{-\frac{y^2}{2p}} dy \Rightarrow p! \approx \sqrt{pe^{p\ln(p)-p}} \int_{-\sqrt{p}}^{\infty} e^{-\frac{y^2}{2p}} dy. \tag{11.62}
\]

Noting that
\[
  \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \sqrt{2\pi} \tag{11.63}
\]
we may write
\[
  \sqrt{2\pi} = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \int_{-\infty}^{\sqrt{p}} e^{-\frac{y^2}{2}} dy + \int_{\sqrt{p}}^{\infty} e^{-\frac{y^2}{2}} dy
\]
\[
  \Rightarrow \int_{-\sqrt{p}}^{\infty} e^{-\frac{y^2}{2}} dy = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy - \int_{-\infty}^{\sqrt{p}} e^{-\frac{y^2}{2}} dy = \sqrt{2\pi} - \int_{-\infty}^{\sqrt{p}} e^{-\frac{y^2}{2}} dy, \tag{11.64}
\]
we can write
\[
  p! \approx \sqrt{pe^{p\ln(p)-p}} \int_{-\sqrt{p}}^{\infty} e^{-\frac{y^2}{2p}} dy = \sqrt{2p\pi e^{p\ln(p)-p}} - \sqrt{pe^{p\ln(p)-p}} \int_{-\infty}^{\sqrt{p}} e^{-\frac{y^2}{2}} dy. \tag{11.65}
\]
11.3. **STIRLING’S FORMULA**

The second integral
\[
\lim_{p \to \infty} \int_{-\infty}^{\sqrt{p}} e^{-y^2/2} dy \to 0.
\]  
(11.66)

Hence, the *Stirling’s approximation* for the factorial function will be

\[ p! \simeq \sqrt{2\pi p}e^{\ln(p)-p} = \sqrt{2\pi p}e^{\ln(p^p)}e^{-p} \simeq \sqrt{2\pi p}p^pe^{-p}. \]  
(11.67)

or with further approximation

\[ \ln(p!) \simeq \ln \left( \sqrt{2\pi p}p^pe^{-p} \right) = \frac{1}{2} \left( \ln \pi + \ln p + p \ln p - p \right) \Rightarrow \ln(p!) \simeq p \ln p - p. \]  
(11.68)

where we dropped the first two terms as compared to the last two terms.

**Example 11.4** Consider a classroom full of gas molecules. There are approximately \( N = 5000N_A = 3 \times 10^{27} \) molecules in the room. From the Binomial Theorem, it can be shown that the probability for \( n \) of the molecules to be in the front half and \( n' = N - n \) molecules to be in the back half of the room is given by

\[ P(n) = \binom{N}{n} p^n q^{n'} = \frac{N!}{n!(N-n)!} p^n q^{N-n}, \]  
(11.69)

where \( p \) is the probability that a molecule will be found in the front half of the room, and \( q \) is the probability that it will be found in the back half. From the symmetry of the problem, we must have

\[ p = \frac{1}{2}, q = 1 - p = \frac{1}{2} = p \]  
(11.70)

On the average, we would expect to find half of the molecules in the front half of the room and the other half of the molecules in the back half of the room. Find the probability that 0.1% of the molecules in the room have shifted from the front to the back half of the room. That is, find the value of \( P(n) = P(0.499N) \).

**Solution:** Since \( q = p \), we can write

\[ P(n) = \frac{N!}{n!(N-n)!} p^n q^{N-n} = \frac{N!}{n!(N-n)!} p^n. \]  
(11.71)

On the average there are \( n_{ave} = 0.5N \) of molecules in the front half of the room and \( n'_{ave} = N - n_{ave} = 0.5N \) in the back half of the room. Here we want to find the probability that 0.1% of the molecules shifted to the back half of the room. In other words, we want to determine the probability that the number of molecules in the front half \( (n_{ave}) \) is reduced by 0.1%. Which means we want to find \( P(n) \) for

\[ n = 0.5N - (N \times 0.1\%) \Rightarrow n = 0.499N \]  
(11.72)
which is given by

\[ P(n) = \frac{N!}{n!(N-n)!}p^n. \]  

(11.73)

Obviously, both \( N \) and \( n \) are very large number and we can make Stirling’s approximation

\[ \ln n! \approx n \ln n - n \]  

(11.74)

for the factorial. Thus

\[ \ln [P(n)] = \ln \left[ \frac{N!}{n!(N-n)!}p^n \right] = \ln N! - \ln n! - \ln (N-n)! + np^N \]  

(11.75)

can be approximated as

\[ \ln [P(n)] \approx N \ln N - N - (n \ln n - n) - ((N-n) \ln (N-n) - (N-n)) + N \ln p = N \ln N - N - n \ln n + n - N \ln (N-n) + n \ln (N-n) + N - n + N \ln p = N [\ln N - \ln (N-n) + \ln p] + n [\ln (N-n) - \ln n] \]

\[ = \frac{N}{N-n} + n \ln \left( \frac{N-n}{n} \right) \]

\[ \Rightarrow \ln [P(n)] \approx n \ln \left( \frac{N-n}{n} \right) - N \ln \left( \frac{N-n}{Np} \right). \]  

(11.76)

Substituting the values

\[ N = 3 \times 10^{27}, n = 0.499N \Rightarrow N - n = 0.501N, p = 0.5, \]  

(11.77)

we find

\[ \ln[6] = n \log \left[ \frac{N}{n} \right] - \log \left[ \frac{N-n}{0.5N} \right] \]

\[ \text{Out}[6] = -6 \times 10^{11} \]

and

\[ P(n) \approx \exp[-6 \times 10^{21}]. \]  

(11.78)

### 11.4 The Error Function

The error function is defined as the area under

\[ \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \]  

(11.79)

There are also other related integrals which sometimes referred as error functions. These includes
11.4. THE ERROR FUNCTION

(a) The standard normal or Gaussian cumulative distribution function, \( \Phi(x) \),

\[
\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt = \frac{1}{2} + \frac{1}{2} \text{erf} \left( \frac{x}{\sqrt{2}} \right). \quad (11.80)
\]

Noting that

\[
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-t^2/2} dt + \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-t^2/2} dt
\]

\[
= \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-t^2/2} dt + \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\pi}{2}} + \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-t^2/2} dt
\]

\[
\Rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-t^2/2} dt \quad (11.81)
\]

The error function can also be expressed as

\[
\frac{1}{2} \text{erf} \left( \frac{x}{\sqrt{2}} \right) = \Phi(x) - \frac{1}{2} = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-t^2/2} dt. \quad (11.82)
\]

(b) The complementary error function:

\[
\text{erf} c(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-t^2/2} dt = 1 - \text{erf} \left( \frac{x}{\sqrt{2}} \right),
\]

\[
\Rightarrow \text{erf} c \left( \frac{x}{\sqrt{2}} \right) = \sqrt{\frac{2}{\pi}} \int_{x}^{\infty} e^{-t^2/2} dt \quad (11.83)
\]

Example 11.3 Consider a criterion that either is or is not satisfied. We look at a system that has many elements, each of which satisfies or does not satisfy the criterion.

For example: Consider a test with many multiple-choice questions. Criterion: the answer to a test question is correct. Each answer on the test is either correct (satisfies criterion) or is incorrect (does not satisfy the criterion).

We then look at a large number of these systems (for example, a large number of tests consisting of multiple-choice questions). We let \( x \) represent the number of elements within a given system satisfying the criterion. We then define the following:

\( \bar{x} \equiv \) The average number of elements satisfying the criterion

\( \sigma \equiv \) The standard deviation about the mean of the number of elements satisfying the criterion.

The probability that any one system will have \( x \) to \( x + dx \) elements satisfying the criterion is then given by the Gaussian distribution:

\[
\alpha(x)dx = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} dx
\]
Find an expression in terms of the error function for the probability that the number of elements of a given system satisfying the criterion, \( x \), will be in the range

\[ \bar{x} - n\sigma \leq x \leq \bar{x} + n\sigma \]

for some real value of \( n \) (usually integral).

**Solution:** The probability that one system will have \( x \) to \( x + dx \) elements satisfying the criterion is

\[ \alpha(x)dx = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} dx \quad (11.84) \]

then the probability that the number of elements in the range \( \bar{x} - n\sigma \leq x \leq \bar{x} + n\sigma \) satisfying the criterion will be

\[ P_n(\bar{x}) = \int_{\bar{x}-n\sigma}^{\bar{x}+n\sigma} \alpha(x)dx = \int_{\bar{x}-n\sigma}^{\bar{x}+n\sigma} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} dx. \quad (11.85) \]

Introducing a new variable defined by

\[ \frac{x-\bar{x}}{\sqrt{2}\sigma} = y \Rightarrow dx = \sqrt{2}\sigma dy \quad (11.86) \]

and noting that for \( x_1 = \bar{x} - n\sigma \) and \( x_2 = \bar{x} + n\sigma \)

\[ y_1 = \frac{x_1 - \bar{x}}{\sqrt{2}\sigma} = \frac{\bar{x} - n\sigma - \bar{x}}{\sqrt{2}\sigma} = -\frac{n}{\sqrt{2}} \]

\[ y_2 = \frac{x_2 - \bar{x}}{\sqrt{2}\sigma} = \frac{\bar{x} + n\sigma - \bar{x}}{\sqrt{2}\sigma} = \frac{n}{\sqrt{2}} \quad (11.87) \]
we can write
\[
P_n(x) = \int_{-\frac{n}{\sqrt{2}}}^{\frac{n}{\sqrt{2}}} \frac{1}{\sqrt{2\pi}\sigma} e^{-y^2} \sqrt{2}\sigma dy = \frac{1}{\sqrt{\pi}} \int_{-\frac{n}{\sqrt{2}}}^{\frac{n}{\sqrt{2}}} e^{-y^2} dy. \tag{11.88}
\]

If we split the integral into two regions \(\left(-\frac{n}{\sqrt{2}}, 0\right)\) and \(\left(0, \frac{n}{\sqrt{2}}\right)\), we have
\[
P_n(x) = \frac{1}{\sqrt{\pi}} \left[ \int_{-\frac{n}{\sqrt{2}}}^{0} e^{-y^2} dy + \int_{0}^{\frac{n}{\sqrt{2}}} e^{-y^2} dy \right] \tag{11.89}
\]
and noting that
\[
\int_{-\frac{n}{\sqrt{2}}}^{0} e^{-y^2} dy \int_{0}^{\frac{n}{\sqrt{2}}} e^{-y^2} dy = \int_{-\frac{n}{\sqrt{2}}}^{\frac{n}{\sqrt{2}}} e^{-y^2} dy \tag{11.90}
\]
we find
\[
P_n(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{\frac{n}{\sqrt{2}}} e^{-y^2} dy. \tag{11.91}
\]
Recalling the definition for the error function
\[
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt \tag{11.92}
\]
we find that
\[
P_n(x) = \text{erf} \left( \frac{n}{\sqrt{2}} \right). \tag{11.93}
\]
You can get the values of the error function for different values of \(n\) using, for example, Mathematica. You will find the following results

\[
\text{In}[3]:= \frac{2}{\sqrt{\pi}} \int_{0}^{\frac{1}{\sqrt{2}}} \exp[-y^2] dy \\
\text{Out}[3]= 0.682689
\]

\[
\text{In}[4]:= \frac{2}{\sqrt{\pi}} \int_{0}^{\frac{2}{\sqrt{2}}} \exp[-y^2] dy \\
\text{Out}[4]= 0.9545
\]

\[
\text{In}[5]:= \frac{2}{\sqrt{\pi}} \int_{0}^{\frac{3}{\sqrt{2}}} \exp[-y^2] dy \\
\text{Out}[5]= 0.9973
\]

<table>
<thead>
<tr>
<th>(n)</th>
<th>(P_n(\pi))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>68.26%</td>
</tr>
<tr>
<td>2</td>
<td>95.44%</td>
</tr>
<tr>
<td>3</td>
<td>99.74%</td>
</tr>
</tbody>
</table>
11.5 Elliptic Integrals

The Complete Elliptic Integral of the First Kind

\[ K(\xi) = \int_{0}^{\pi/2} \frac{d\varphi}{\sqrt{1 - \xi^2 \sin^2 \varphi}}. \]  

(11.94)

Example 11.5 Consider a simple pendulum with a mass \( m \) suspended from the end of a light rigid rod of length \( l \). We pull the pendulum to the side by an angle \( \alpha \) and release it from rest. Find an expression for the period of the pendulum, \( T \), assuming that \( \dot{\theta} = 0 \) and \( d\theta/dt > 0 \) at \( t = 0 \), where \( \theta \) is the angle of the pendulum from the vertical. Then find an approximate expression for the period of the pendulum for not-so-small amplitudes of motion.

Solution: Using conservation of Mechanical energy,

\[ ME_I = ME_\theta \]  

(11.95)

where \( ME_I \) is the initial mechanical energy (when the pendulum is pulled to the side by an angle \( \alpha \)) which is just only the gravitational potential energy given by

\[ ME_I = mgh_{\text{max}} = mgl(1 - \cos \alpha) \]  

(11.96)

Figure 11.2: A simple pendulum. At the initial time, \( t = 0 \), the mass \( m \) was displaced by an angle, \( \alpha \), from the vertical.
and $ME_\theta$ is the mechanical energy at some instant of time (i.e. at an angle $\theta$) which is the sum of kinetic and potential energy given by

$$ME_\theta = mgh + \frac{1}{2}mv^2 = mgl(1 - \cos \theta) + \frac{1}{2}ml^2 \left( \frac{d\theta}{dt} \right)^2.$$  \hspace{1cm} (11.97)

Then

$$ME_I = ME_\theta \Rightarrow mgl(1 - \cos \alpha) = mgl(1 - \cos \theta) + \frac{1}{2}ml^2 \left( \frac{d\theta}{dt} \right)^2$$

$$\Rightarrow -g \cos \alpha = -g \cos \theta + \frac{1}{2}l \left( \frac{d\theta}{dt} \right)^2 \Rightarrow \frac{d\theta}{dt} = \sqrt{\frac{2g}{l} (\cos \theta - \cos \alpha)}.$$  \hspace{1cm} (11.98)

Noting that the period is the time for one complete oscillation which we can express as

$$T = 2 \int_{0}^{t_{1/2}} dt$$  \hspace{1cm} (11.99)

where $t = 0$ is the time at which the pendulum at the maximum displacement from the vertical ($\theta = -\alpha$) and $t = t_{1/2}$ is the time at which the pendulum reached to the position ($\theta = \alpha$). Therefore, using

$$\frac{d\theta}{dt} = \sqrt{\frac{2g}{l} (\cos \theta - \cos \alpha)} \Rightarrow dt = \frac{d\theta}{\sqrt{2g/l} (\cos \theta - \cos \alpha)}$$

we can write

$$T = 2 \int_{0}^{t_{1/2}} dt = 2 \int_{-\alpha}^{\alpha} \frac{d\theta}{\sqrt{2g/l} (\cos \theta - \cos \alpha)}$$  \hspace{1cm} (11.100)

Using the

$$\cos \theta = 1 - 2 \sin^2 \left( \frac{\theta}{2} \right), \cos \alpha = 1 - 2 \sin^2 \left( \frac{\alpha}{2} \right)$$  \hspace{1cm} (11.102)

we have

$$\sqrt{\frac{2g}{l} (\cos \theta - \cos \alpha)} = \sqrt{\frac{2g}{l} \left( 1 - 2 \sin^2 \left( \frac{\theta}{2} \right) - 1 + 2 \sin^2 \left( \frac{\alpha}{2} \right) \right)}$$

$$= \sqrt{\frac{2g}{l} \left( 2 \sin^2 \left( \frac{\alpha}{2} \right) - 2 \sin^2 \left( \frac{\theta}{2} \right) \right)}$$  \hspace{1cm} (11.103)

so that

$$T = 2 \int_{-\alpha}^{\alpha} \frac{d\theta}{\sqrt{\frac{2g}{l} \left( 2 \sin^2 \left( \frac{\alpha}{2} \right) - 2 \sin^2 \left( \frac{\theta}{2} \right) \right)}}$$

$$= \sqrt{\frac{l}{g}} \int_{-\alpha}^{\alpha} \frac{d\theta}{\sqrt{\sin^2 \left( \frac{\alpha}{2} \right) - \sin^2 \left( \frac{\theta}{2} \right)}} = \sqrt{\frac{l}{g}} \int_{-\alpha}^{\alpha} \frac{d\theta}{\sqrt{1 - \sin^2 \left( \frac{\theta}{2} \right)}}$$  \hspace{1cm} (11.104)
Introducing the transformation

\[ \sin \varphi = \sin \left( \frac{\theta}{2} \right) \Rightarrow \sqrt{1 - \sin^2 \left( \frac{\theta}{2} \right)} = \cos \varphi \]

\[ \Rightarrow \cos \varphi d\varphi = \frac{1}{2} \cos \left( \frac{\theta}{2} \right) d\theta \Rightarrow \frac{d\theta}{\sin \left( \frac{\theta}{2} \right)} = 2 \cos \varphi \frac{d\varphi}{\cos \left( \frac{\theta}{2} \right)} = \frac{2 \cos \varphi}{\sqrt{1 - \sin^2 \left( \frac{\theta}{2} \right)}} \]

\[ \Rightarrow \frac{d\theta}{\sin \left( \frac{\theta}{2} \right)} = \frac{2 \cos \varphi}{\sqrt{1 - \sin^2 \left( \frac{\theta}{2} \right) \sin^2 \varphi}}, \tag{11.105} \]

\[ \theta = \pm \alpha \Rightarrow \sin \varphi = \pm 1 \Rightarrow \varphi = \pm \frac{\pi}{2} \tag{11.106} \]

The expression for the period can be put in the form

\[ T = 2 \sqrt{\frac{l}{g}} \int_{-\pi/2}^{\pi/2} \frac{d\varphi}{\sqrt{1 - \sin^2 \left( \frac{\theta}{2} \right) \sin^2 \varphi}} = 4 \sqrt{\frac{l}{g}} \int_{0}^{\pi/2} \frac{d\varphi}{\sqrt{1 - \xi^2 \sin^2 \varphi}}, \]

where we introduced the constant

\[ \xi = \sin \left( \frac{\alpha}{2} \right) \tag{11.107} \]

and take into consideration the fact that \( \sin^2 \varphi \) is an even function such that

\[ \int_{-\pi/2}^{\pi/2} \frac{d\varphi}{\sqrt{1 - \sin^2 \left( \frac{\theta}{2} \right) \sin^2 \varphi}} = 2 \int_{0}^{\pi/2} \frac{d\varphi}{\sqrt{1 - \sin^2 \left( \frac{\theta}{2} \right) \sin^2 \varphi}}. \]

Therefore, the period is given by

\[ T = 4 \sqrt{\frac{l}{g}} K (\xi) \tag{11.108} \]

where

\[ K (\xi) = \int_{0}^{\pi/2} \frac{d\varphi}{\sqrt{1 - \xi^2 \sin^2 \varphi}} \tag{11.109} \]

is the elliptic integral of the first kind.

### 11.6 The Dirac delta function

Let’s consider a point in space that can be described by the Cartesian coordinates \((x, y, z)\) or the spherical coordinates \((r, \theta, \varphi)\) as shown in Fig. The position vector

\[ \vec{r} = x \hat{x} + y \hat{y} + z \hat{z} = r \sin (\theta) \cos (\varphi) \hat{x} + r \sin (\theta) \sin (\varphi) \hat{y} + r \cos (\theta) \hat{z} = r \hat{r} \tag{11.110} \]
11.6. THE DIRAC DELTA FUNCTION

$$\hat{r} = \sin (\theta) \cos (\varphi) \hat{x} + \sin (\theta) \sin (\varphi) \hat{y} + \cos (\theta) \hat{z} = \frac{\partial r}{\partial r}$$ (11.111)

is the unit vector along the radial direction. In spherical coordinates we recall the gradient and the Laplacian for a scalar function, \( f (r, \theta, \varphi) \), are given by

$$\nabla f = r \frac{\partial f}{\partial r} + \theta \frac{1}{r} \frac{\partial f}{\partial \theta} + \varphi \frac{1}{r \sin (\theta)} \frac{\partial f}{\partial \varphi},$$ (11.112)

and

$$\nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin (\theta)} \frac{\partial}{\partial \theta} \left( \sin (\theta) \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 (\theta)} \frac{\partial^2 f}{\partial \varphi^2},$$ (11.113)

respectively. On the other hand for a vector field,

$$\vec{V}(r, \varphi, \theta) = V_r(r, \varphi, \theta) \hat{r} + V_{\theta}(r, \varphi, \theta) \hat{\theta} + V_{\varphi}(r, \varphi, \theta) \hat{\varphi},$$ (11.114)

instead of the gradient, most often, we are interested in the divergence of the vector field. In spherical coordinates the divergence of a vector field is given by

$$\nabla \cdot \vec{V} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 V_r \right) + \frac{1}{r \sin (\theta)} \frac{\partial}{\partial \theta} \left( \sin (\theta) V_{\theta} \right) + \frac{1}{r \sin (\theta)} \frac{\partial V_{\varphi}}{\partial \varphi}.$$ (11.115)
Suppose we have some physical quantity described by some scalar function that
depend on the radial distance between two points. A good example of such phys-
ical quantity is electrical and gravitational potentials. Let’s say this function in
spherical coordinates can be expressed as

\[ f(r, \theta, \varphi) = \frac{1}{r}, \quad \text{for } r > 0, \quad (11.116) \]

and one can construct a vector field

\[ \vec{V} = \nabla \left( \frac{1}{r} \right) = \hat{r} \frac{\partial}{\partial r} \left( \frac{1}{r} \right) = -\frac{\hat{r}}{r^2}, \quad \text{for } r > 0. \quad (11.117) \]

The divergence of this vector field becomes

\[ \nabla \cdot \vec{V} = \nabla \cdot \left( -\frac{\hat{r}}{r^2} \right) = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \left( -\frac{1}{r^2} \right) \right) = \begin{cases} 0, & \text{for } r > 0 \\ \infty, & \text{for } r = 0 \end{cases}. \quad (11.118) \]

There follows that

\[ \nabla^2 \left( \frac{1}{r} \right) = \nabla \cdot \nabla \left( \frac{1}{r} \right) = \nabla \cdot \left( -\frac{\hat{r}}{r^2} \right) = \begin{cases} 0, & \text{for } r > 0 \\ \infty, & \text{for } r = 0 \end{cases}, \quad (11.119) \]

which leads to

\[ \iiint_{\text{volume}} \left( \nabla \cdot \vec{V} \right) \, d\tau = \iiint_{\text{volume}} \nabla \cdot \left( -\frac{\hat{r}}{r^2} \right) \, d\tau = 0, \quad \text{for } r > 0 \quad (11.120) \]

\[ \Rightarrow \int_{0}^{\infty} \int_{0}^{2\pi} \int_{0}^{\pi} \nabla \cdot \left( -\frac{\hat{r}}{r^2} \right) \, d\tau = 0, \quad \text{for } r > 0 \quad (11.121) \]

On the other hand if one determines using the divergence theorem

\[ \iiint_{\text{volume}} \left( \nabla \cdot \vec{V} \right) \, d\tau = \iint_{\text{Surface \, enclosing}} \vec{V}(x, y, z) \cdot d\vec{a} \quad (11.122) \]

we find

\[ \iint_{\text{Surface \, enclosing}} \vec{V}(x, y, z) \cdot d\vec{a} \]

\[ = \int_{0}^{\infty} \int_{0}^{\frac{\pi}{2}} \int_{0}^{\frac{2\pi}{2}} \left( -\frac{\hat{r}}{r^2} \right) \cdot \hat{r}r^2 \sin(\theta) \cos(\varphi) \, d\theta d\varphi = -\pi \int_{0}^{\frac{\pi}{2}} \sin(\theta) \cos(\varphi) \, d\theta d\varphi = -4\pi \]

\[ = \int_{0}^{\frac{\pi}{2}} \int_{0}^{\frac{2\pi}{2}} \left( -\frac{\hat{r}}{r^2} \right) \cdot d\vec{a} \]

(11.123)

This none zero value must be a result when \( r = 0 \) is included. Therefore, one
can write

\[ \iiint_{0}^{\infty} \int_{0}^{\frac{\pi}{2}} \int_{0}^{\frac{2\pi}{2}} \nabla \cdot \left( -\frac{\hat{r}}{r^2} \right) \, d\tau = \begin{cases} 0, & \text{for } r > 0 \\ -4\pi, & \text{for } r = 0 \end{cases} \quad (11.124) \]
11.6. THE DIRAC DELTA FUNCTION

The function, \( \sigma (r) \), defined as

\[
4 \pi \sigma (r) = \nabla \cdot \left( \frac{\hat{r}}{r^2} \right) = -\nabla \cdot \nabla \left( \frac{1}{r} \right) = -\nabla^2 \left( \frac{1}{r} \right) = \begin{cases} 0 & \text{for } r \neq 0 \\ \infty & \text{for } r = 0 \end{cases}
\]

is known as the Dirac delta function. Then for the Dirac delta function

\[
\int_0^\infty \int_0^{2\pi} \int_0^{2\pi} \sigma (r) \, d\tau = 1, \\
\Rightarrow f (0) \int_0^\infty \int_0^{2\pi} \int_0^{2\pi} \sigma (r) \, d\tau = f (0) \Rightarrow \int_0^\infty \int_0^{2\pi} f (\hat{r}) \sigma (r) \, d\tau = f (0),
\]

(11.126)

For one dimensional case,

\[
\sigma (x) = \begin{cases} 0 & \text{for } x \neq 0 \\ \infty & \text{for } x = 0 \end{cases}
\]

(11.127)

and

\[
\int_{-\infty}^{\infty} \sigma (x) \, dx = 1, \quad \int_{-\infty}^{\infty} f (x) \sigma (x) \, dx = f (0)
\]

(11.128)

In the usual sense of functions the Dirac delta function does not exist. But there are various forms of sequence functions in the limiting case display the properties of the Dirac delta function. These functions include:

1. The Gaussian function:

\[
\sigma_n (x) = \frac{n}{\sqrt{\pi}} \exp \left( -n^2 x^2 \right),
\]

See Fig. 1.

2. Lorentz function

\[
\sigma_n (x) = \frac{n}{\pi} \frac{1}{1 + n^2 x^2},
\]

See Fig. 2.

3. The sinc function:

\[
\sigma_n (x) = \frac{\sin (nx)}{nx} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixt} \, dt
\]

See Fig. ?? . The general form of the Dirac delta function for, \( x = a \)

\[
\sigma (x - a) = \begin{cases} 0 & \text{for } x \neq a \\ \infty & \text{for } x = a \end{cases}
\]
and

\[ \int_{-\infty}^{\infty} \sigma(x-a) \, dx = 1, \quad \int_{-\infty}^{\infty} f(x) \sigma(x-a) \, dx = f(a). \]  \hspace{1cm} (11.130)
For 3D case

\[
\int \int \int f(r) \sigma(r) r^2 \sin(\theta) \, dr \, d\theta \, d\varphi = f(0)
\]

\[
\int \int \int f(r) \sigma(r) r^2 \sin(\theta) \, dr \, d\theta \, d\varphi = f(\mathbf{r}_0) \quad (11.131)
\]

**Example 11.6** Form introductory physics, the electric potential, \( V(\mathbf{r}) \), due to a point charge located at the origin \((0, 0, 0)\) (i.e. \( r = 0 \)) at a point in space described by the position vector, \( \mathbf{r} \), is given by

\[
V(\mathbf{r}) = \frac{1}{4\pi \varepsilon_0} \frac{q}{r}.
\]

Show that the volume charge density, \( \rho(\mathbf{r}) \), for this point charge can be expressed in terms of the Dirac delta function

\[
\rho(\mathbf{r}) = \frac{dq}{d\tau} = \sigma(\mathbf{r}) = q\sigma(x)\sigma(y)\sigma(z),
\]

where \( dq \) an infinitesimal charge in an infinitesimal volume \( d\tau \).

**Solution:** The electric potential, \( dV(\mathbf{r}) \) of an infinitesimal charge \( dq' \) in a volume \( d\tau' \) as shown in Fig. can be expressed as

\[
dV(\mathbf{r}) = \frac{1}{4\pi \varepsilon_0} \frac{dq'}{|\mathbf{r}' - \mathbf{r}|} = \frac{1}{4\pi \varepsilon_0} \rho(\mathbf{r}') \, d\tau' \Rightarrow V(\mathbf{r}) = \frac{1}{4\pi \varepsilon_0} \int \int V \rho(\mathbf{r}') \, d\tau' \]

\[
\quad = \frac{1}{4\pi \varepsilon_0} \int \int \frac{\rho(\mathbf{r}') \, d\tau'}{|\mathbf{r}' - \mathbf{r}|} \quad (11.133)
\]
Using spherical coordinates, we can write
\[
V(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \int_0^\infty \int_0^{2\pi} \int_0^\pi \frac{\rho(\vec{r}') r'^2 \sin(\theta)' \, d\theta' \, d\phi'}{|\vec{r}' - \vec{r}|} \quad (11.134)
\]
This potential for a point charge becomes
\[
\frac{1}{4\pi\varepsilon_0} \int_0^\infty \int_0^{2\pi} \int_0^\pi \frac{\rho(\vec{r}') r'^2 \sin(\theta)' \, d\theta' \, d\phi'}{|\vec{r}' - \vec{r}|} = \frac{1}{4\pi\varepsilon_0} \frac{q}{r}
\]
\[
\Rightarrow \int_0^\infty \int_0^{2\pi} \int_0^\pi \frac{\rho(\vec{r}') r'^2 \sin(\theta)' \, d\theta' \, d\phi'}{|\vec{r}' - \vec{r}|} = \frac{1}{r}. \quad (11.135)
\]
From the property of the Dirac Delta function
\[
\int_0^\infty \int_0^{2\pi} \int_0^\pi f(\vec{r}') \sigma(\vec{r}' - \vec{r}_0) r'^2 \sin(\theta)' \, d\theta' \, d\phi' = f(\vec{r}_0) \quad (11.136)
\]
one can easily find
\[
f(\vec{r}') = \frac{1}{|\vec{r}' - \vec{r}|} \sigma(\vec{r}' - \vec{r}_0) = \frac{\rho(\vec{r}')}{q}
\]
\[
\Rightarrow f(\vec{r}_0) = \frac{1}{|\vec{r} - \vec{r}_0|} = \frac{1}{r} \Rightarrow \vec{r}_0 = 0
\]
which leads to
\[
\frac{\rho(\vec{r}')}{q} = \sigma(\vec{r}') \Rightarrow \rho(\vec{r}') = q \sigma(\vec{r}') \quad (11.137)
\]
**Example 11.7** The volume charge density, \(\rho(\vec{r})\), of a point charge, \(q\), placed at a point on the z-axis, \(\vec{r}_0 = a \hat{z}\), can be expressed as
\[
\rho(\vec{r}) = q \sigma(\vec{r}' - \vec{r}_0), \quad (11.139)
\]
where \(\sigma(\vec{r})\) is the Dirac Delta function. Show that the electric potential, \(V(\vec{r})\), due to this point charge is given by
\[
V(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \frac{q}{|\vec{r} - \vec{r}_0|} = \frac{1}{4\pi\varepsilon_0} \frac{q}{\sqrt{x'^2 + y'^2 + (z - a)^2}} \quad (11.140)
\]
The electric potential for a volume charge distribution is given by
\[
V(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \iiint_V \frac{\rho(\vec{r}') \, d\tau'}{|\vec{r} - \vec{r}'|}, \quad (11.141)
\]
where \(\vec{r}'\) is the position of the infinitesimal charge \(dq' = \rho(\vec{r}') \, d\tau'\), in an infinitesimal volume \(d\tau'\), \(\rho(\vec{r}')\) is the charge density in the volume \(V\).
Solution: Using the given charge density and the expression for the potential, one can write

\[ V(\vec{r}) = \frac{q}{4\pi \varepsilon_0} \int \frac{\sigma(\vec{r} - \vec{r}_0) d\vec{r}'}{|\vec{r} - \vec{r}'|} = \frac{q}{4\pi \varepsilon_0} \int \frac{\sigma(\vec{r} - \vec{r}_0) d\vec{r}'}{|\vec{r} - \vec{r}'|}, \quad (11.142) \]

In Cartesian coordinates, we have

\[ \vec{r}' = x'\hat{x} + y'\hat{y} + z'\hat{z}, \vec{r} = x\hat{x} + y\hat{y} + z\hat{z}, \vec{r}_0 = a\hat{z} \]

\[ \Rightarrow |\vec{r} - \vec{r}'| = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2} \quad (11.143) \]

and

\[ \sigma(\vec{r} - \vec{r}_0) = \sigma(x' - x_0) \sigma(y' - y_0) \sigma(z' - z_0) \]

\[ = \sigma(x') \sigma(y') \sigma(z' - a) \quad (11.144) \]

so that

\[ V(\vec{r}) = \frac{q}{4\pi \varepsilon_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\sigma(x') \sigma(y') \sigma(z' - a) dx' dy' dz'}{\sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}} \]

\[ = \frac{q}{4\pi \varepsilon_0} \int_{-\infty}^{\infty} \sigma(z' - a) dz' \int_{-\infty}^{\infty} \sigma(y') dy' \int_{-\infty}^{\infty} f(x', y', z') \sigma(x)(x' dx') \]

where

\[ f(x', y', z') = \frac{1}{\sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}}. \quad (11.146) \]

Now applying the property of the Dirac Delta function

\[ \int_{-\infty}^{\infty} f(x) \sigma(x - a) dx = f(a) \quad (11.147) \]

one can easily see that

\[ \int_{-\infty}^{\infty} f(x', y', z') \sigma(x') dx' = \int_{-\infty}^{\infty} f(x', y', z') \sigma(x' - 0) dx' \]

\[ = f(0, y', z') = \frac{1}{\sqrt{x^2 + (y - y')^2 + (z - z')^2}} \quad (11.148) \]

The electric potential becomes

\[ V(\vec{r}) = \frac{q}{4\pi \varepsilon_0} \int_{-\infty}^{\infty} \sigma(z' - a) dz' \int_{-\infty}^{\infty} f(0, y', z') \sigma(y') dy'. \quad (11.149) \]
Once again using the property of the Dirac delta function, we have

\[
\int_{-\infty}^{\infty} f(0, y', z') \sigma(y') \, dy' = \int_{-\infty}^{\infty} f(0, y', z') \sigma(y' - 0) \, dy' \\
= f(0, 0, z') = \frac{1}{\sqrt{x^2 + y^2 + (z - z')^2}}, \quad (11.150)
\]

and the expression for potential reduces to

\[
V(\vec{r}) = \frac{q}{4\pi \varepsilon_0} \int_{-\infty}^{\infty} f(0, 0, z') \sigma(z' - a) \, dz'. \quad (11.151)
\]

One last time using the Dirac delta function property, we find for the potential

\[
V(\vec{r}) = \frac{q}{4\pi \varepsilon_0} \int_{-\infty}^{\infty} f(0, 0, z') \sigma(z' - a) \, dz' = \frac{q}{4\pi \varepsilon_0} f(0, 0, a) \\
\Rightarrow V(\vec{r}) = \frac{q}{4\pi \varepsilon_0} \frac{1}{\sqrt{x^2 + y^2 + (z - a)^2}}. \quad (11.152)
\]
Chapter 12

Power Series Solutions to Differential Equations

12.1 Power series substitution

The differential equations we seek to find the solution to are linear differential equation like those we studied in PHYS 3150. However, unlike those equations here the coefficients in the differential equations we shall consider are not constants and depend on the variable, \( x \), like the following differential equation

\[
\frac{d^2y(x)}{dx^2} + f(x) \frac{dy(x)}{dx} + g(x)y(x) = 0. \tag{12.1}
\]

Such kind of differential equations or even those with constant coefficients can be solved using series substitution method. The method involves a simple procedure but a little too much algebra. We assume the solution to the differential equation can be expressed as a power series

\[
y(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots a_n x^n + \ldots \tag{12.2}
\]

We then substitute this series into the differential equation and determine the expansion coefficients. We will demonstrate the application of this method using the following example which we already determined the solution using a different method last semester.

Example 12.1 A mass \( m \) is attached to a horizontal spring of spring constant, \( k \), whose other end is attached to a rigid vertical wall. The mass slides on a horizontal, frictionless surface. At time, \( t = 0 \), the mass is stretched from its equilibrium position by a distance \( x = x_{\text{max}} \) and released from rest. Find the equation for the position of the mass as a function of time, \( x(t) \).
Solution: Using Newton’s second law the equation of motion for the mass, \( m \), can be written as

\[
F_{\text{net}} = ma \Rightarrow m \frac{d^2x}{dt^2} = -kx, \quad (12.3)
\]

which we may rewrite as

\[
\frac{d^2x}{dt^2} + \omega^2 x = 0, \quad (12.4)
\]

where

\[
\omega = \sqrt{\frac{k}{m}}, \quad (12.5)
\]

is the angular frequency. Let’s assume that the solution to this differential equation is given by the power series

\[
x(t) = \sum_{n=0}^{\infty} a_n t^n = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \ldots a_n t^n + \ldots \quad (12.6)
\]

so that

\[
\frac{dx(t)}{dt} = \sum_{n=0}^{\infty} a_n \frac{d(t^n)}{dt} = \sum_{n=0}^{\infty} n a_n t^{n-1} = a_1 + 2a_2 t + 3a_3 t^2 + \ldots \]

\[
\ldots n a_n t^{n-1} + \ldots = \sum_{n=0}^{\infty} (n+1) a_{n+1} t^n \quad (12.7)
\]

and

\[
\frac{d^2x(t)}{dt^2} = \sum_{n=0}^{\infty} a_n \frac{d^2(t^n)}{dt^2} = \sum_{n=0}^{\infty} n(n-1) a_n t^{n-2}
\]

\[
= 2a_2 + 3 \cdot 2a_3 t + 4 \cdot 3a_4 t^2 + \ldots n(n-1) a_n t^{n-2} + \ldots
\]

\[
= \sum_{n=0}^{\infty} (n+2)(n+1) a_{n+2} t^n \quad (12.8)
\]
The differential equation can then be rewritten as
\[
\sum_{n=0}^{\infty} (n + 2) (n + 1) a_{n+2} t^n + \omega^2 \sum_{n=0}^{\infty} a_n t^n = 0
\]
\[
\Rightarrow \sum_{n=0}^{\infty} [(n + 2) (n + 1) a_{n+2} + \omega^2 a_n] t^n = 0. \quad (12.9)
\]

From the above equation we find the following recursion relation
\[
(n + 2) (n + 1) a_{n+2} + \omega^2 a_n = 0 \Rightarrow a_{n+2} = -\frac{a_n}{(n + 2) (n + 1)} \omega^2.
\]

Now let’s examine the first few terms for this recursion relation. First we consider when \(n\) is even
\[
n = 0 \Rightarrow a_2 = -\frac{a_0}{2 \cdot 1} \omega^2
\]
\[
\Rightarrow n = 0 \Rightarrow a_{2 \times 1} = \frac{a_0 (-1)^1}{2!} \omega^{2 \times 1}
\]
\[
n = 2 \Rightarrow a_4 = -\frac{a_2}{4 \cdot 3} \omega^2 = \frac{a_0 (-1)^2 \omega^{2 \times 2}}{4 \cdot 3 \cdot 2 \cdot 1} = \frac{a_0 (-1)^2 \omega^{2 \times 2}}{4!}
\]
\[
\Rightarrow n = 2 \Rightarrow a_{2 \times 2} = \frac{a_0}{(2 \times 2)!} (-1)^2 \omega^{2 \times 2}
\]
\[
n = 4 \Rightarrow a_6 = -\frac{a_4}{6 \cdot 5} \omega^2 = \frac{a_0 (-1)^3 \omega^{2 \times 3}}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{a_0 (-1)^3 \omega^{2 \times 3}}{6!}
\]
\[
\Rightarrow n = 4 \Rightarrow a_{2 \times 3} = \frac{a_0 (-1)^3 \omega^{2 \times 3}}{(2 \times 3)!}. \quad (12.10)
\]

From the results we see above it is not hard to come up with the relation
\[
a_{2m} = \frac{a_0 (-1)^m \omega^{2m}}{(2m)!} \quad (12.11)
\]

that generates the values for the even term coefficients in the series for \(m = 0, 1, 2, \ldots\)

Next we shall consider the odd terms for the recursion relation
\[
a_{n+2} = -\frac{a_n}{(n + 2) (n + 1)} \omega^2. \quad (12.12)
\]
For the first three odd terms, we have

\[
  n = 1 \Rightarrow a_3 = \frac{a_1 (-1)^1}{3 \cdot 2} \omega^2 = \frac{a_1 (-1)^1}{3!} \cdot \omega^2
\]

\[
  n = 1 \Rightarrow a_{(2 \times 1 + 1)} = a_1 (-1)^1 \omega^{2 \times 1} = \frac{a_1 (-1)^1}{(2 \times 1 + 1)!} \cdot \omega^2
\]

\[
  n = 3 \Rightarrow a_5 = -\frac{a_3 \omega^2}{5 \cdot 4} = \frac{a_1 (-1)^2 \omega^4}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{a_1 (-1)^2 \omega^4}{5!} \cdot \omega^2
\]

\[
  n = 3 \Rightarrow a_{(2 \times 3 - 1)} = a_1 (-1)^2 \omega^{2 \times 2} = \frac{a_1 (-1)^2 \omega^2}{(2 \times 2 + 1)!} \cdot \omega^2
\]

\[
  n = 5 \Rightarrow a_7 = -\frac{a_5 \omega^2}{7 \cdot 6} = \frac{a_1 (-1)^3 \omega^6}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{a_1 (-1)^3 \omega^6}{7!} \cdot \omega^2
\]

\[
  n = 5 \Rightarrow a_{(2 \times 4 - 1)} = \frac{a_1 (-1)^3 \omega^{2 \times 3}}{(2 \times 3 + 1)!} \cdot \omega^2
\]

Thus the general solution to the differential equation

\[
  x(t) = \sum_{n=0}^{\infty} a_n t^n = a_0 \cos(\omega t) + a_1 \sin(\omega t)
\]

using the results obtained above, can be expressed as

\[
  x(t) = \sum_{n=0}^{\infty} a_n t^n = \sum_{m=0}^{\infty} a_0 (-1)^m \omega^{2m} \cdot \omega^{2m} t^{2m} + \sum_{m=0}^{\infty} a_1 (-1)^m \omega^{2m} \cdot \omega^{2m+1} t^{2m+1}
\]

\[
  x(t) = a_0 \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} (\omega t)^{2m} + a_1 \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m + 1)!} (\omega t)^{2m+1}
\]

We recall that the Taylor series expansions for \(\sin(x)\) and \(\cos(x)\) are

\[
  \sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}, \quad \cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}
\]

so that the solution to the differential equation can be put in the form

\[
  x(t) = a_0 \cos(\omega t) + \frac{a_1}{\omega} \sin(\omega t)
\]

or

\[
  x(t) = A \cos(\omega t) + B \sin(\omega t)
\]
where \( A = a_0 \) and \( B = a_1/\omega \) are constants determined by the initial conditions. Since initially \((t = 0)\) the spring is stretched by \( x_{\text{max}} \) and the mass is released from rest, we have

\[
x(t = 0) = a_0 \cos(0) - a_1 \omega \sin(0) = x_{\text{max}} \Rightarrow a_0 = x_{\text{max}}
\]

(12.20)

and

\[
\frac{dx(t)}{dt} = -a_0 \sin(\omega t) - a_1 \omega^2 \cos(\omega t),
\]

\[
\frac{dx(0)}{dt} = 0 \Rightarrow a_1 \omega^2 = 0 \Rightarrow a_1 = 0.
\]

(12.21)

Therefore, the equation for the position of the mass as a function of time, \( x(t) \) is given by

\[
x(t) = x_{\text{max}} \cos(\omega t),
\]

(12.22)

where

\[
\omega = \sqrt{\frac{k}{m}},
\]

(12.23)

is the angular frequency.

### 12.2 Orthogonal vectors and Dirac Notation

We recall that if the scalar product of two real vectors, \( \vec{A} \) and \( \vec{B} \), is zero, given by

\[
\vec{A} \cdot \vec{B} = \sum_{i=1}^{3} A_i B_i = 0,
\]

(12.24)

the two vectors are said to be orthogonal. If these vectors are complex and orthogonal, we must write

\[
\vec{A}^* \cdot \vec{B} = \sum_{i=1}^{3} A_i^* B_i = 0,
\]

(12.25)

where \( \vec{A}^* \) is the complex conjugate of \( \vec{A} \). Using Dirac notation any vector \( \vec{A} \) is denoted by a ket vector \(|\vec{A}\rangle\) and its complex conjugate \( \vec{A}^* \) by a bra vector \( \langle \vec{A}| \). Then the dot product of two vectors, using Dirac notation, can be expressed as

\[
\langle A | B \rangle = \sum_{i=1}^{3} A_i^* B_i
\]

(12.26)

and if the vectors are orthogonal,

\[
\langle A | B \rangle = \sum_{i=1}^{3} A_i^* B_i = 0.
\]

(12.27)
CHAPTER 12. POWER SERIES SOLUTIONS TO DIFFERENTIAL EQUATIONS

Orthonormal Sets of Functions: two different functions \( A(x) \) and \( B(x) \) are said to be orthogonal for all \( x \in (a, b) \) when

\[
\int_a^b A^*(x)B(x)dx = 0, \quad (12.28)
\]

where \( A^*(x) \) is the complex conjugate of the function \( A(x) \). A set of functions

\[
\{A_1(x), A_2(x), A_3(x) \ldots A_n(x) \ldots \} \quad (12.29)
\]

that meet the requirement

\[
\int_a^b A_n^*(x)A_m(x)dx = \begin{cases} 0 & m \neq n, \\ C & m = n, \end{cases} \quad (12.30)
\]

where \( C \) is a constant, are said to form an orthonormal set of functions. Using the Kronecker delta

\[
\delta_{nm} = \begin{cases} 0 & n \neq m \\ 1 & n = m \end{cases}, \quad (12.31)
\]

and Dirac notation, we may write

\[
\langle A_n(x) | A_m(x) \rangle = C_n \delta_{nm} \quad (12.32)
\]

For \( m = n \),

\[
\langle A_n(x) | A_n(x) \rangle = C_n \quad (12.33)
\]

so that the for the set of functions defined by

\[
F_n(x) = \frac{A_n(x)}{\sqrt{C_n}} \quad (12.34)
\]

one can write

\[
\langle F_n(x) | F_m(x) \rangle = \int_a^b F_n^*(x)F_m(x)dx = \delta_{nm}. \quad (12.35)
\]

These set of functions are said to form an orthonormal set of functions.

**Example 12.2** Show that the set of functions defined by

\[
F_n(x) = \frac{1}{\sqrt{\pi}} \sin(nx) \quad (12.36)
\]

for \( n = 1, 2, 3 \ldots \) form an orthonormal set of functions for all \( x \in (-\pi, \pi) \).

**Solution:** Using Euler’s formula, we can express

\[
F_m(x) = \frac{1}{\sqrt{\pi}} \sin(mx) = \frac{e^{imx} - e^{-imx}}{2i\sqrt{\pi}} \quad \Rightarrow \quad F_n^*(x) = \frac{e^{-inx} - e^{inx}}{-2i\sqrt{\pi}} \quad (12.36)
\]
so that

\[ \langle F_n(x) F_m(x) \rangle = \int_a^b F_n^*(x) F_m(x) \, dx \]  

(12.37)

becomes

\[
\langle F_n(x) | F_m(x) \rangle = \int_{-\pi}^{\pi} \left( \frac{e^{inx} - e^{-inx}}{-2i\pi} \right) \left( \frac{e^{imx} - e^{-imx}}{2i\pi} \right) \, dx \\
= \frac{1}{4\pi} \int_{-\pi}^{\pi} \left( e^{i(m-n)x} - e^{-i(m-n)x} \right) \left( e^{i(m+n)x} - e^{-i(m+n)x} \right) \, dx \\
= \frac{1}{4\pi} \left[ \frac{e^{i(m-n)x}}{m-n} + \frac{e^{-i(m+n)x}}{m+n} - \frac{e^{i(m+n)x}}{m+n} - \frac{e^{-i(m-n)x}}{m-n} \right]_{-\pi}^{\pi} \\
\Rightarrow \langle F_n(x) | F_m(x) \rangle = \frac{1}{4\pi} \left( \frac{4\sin((m-n)\pi)}{m-n} - \frac{4\sin((m+n)\pi)}{m+n} \right),
\]  

(12.38)

whether \( m = n \) or \( m \neq n \) the second term is always zero and we can write

\[
\langle F_n(x) | F_m(x) \rangle = \sin \left( \frac{(m-n)\pi}{m-n} \right) = \delta_{nm},
\]

where we have applied applied L’Hospital’s rule. But it would be shorter if one applies the double angle relations

\[
\cos (\alpha + \beta) = \cos (\alpha) \cos (\beta) - \sin (\alpha) \sin (\beta), \\
\cos (\alpha - \beta) = \cos (\alpha) \cos (\beta) + \sin (\alpha) \sin (\beta),
\]

that give

\[
\sin (\alpha) \sin (\beta) = \frac{1}{2} [\cos (\alpha - \beta) - \cos (\alpha + \beta)].
\]

In view of these relation, one can write the integral as

\[
\langle F_n(x) | F_m(x) \rangle = \int_{-\pi}^{\pi} \sin (nx) \sin (mx) \, dx \\
= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[ \cos ((n-m)x) - \cos ((n+m)x) \right] \, dx \\
= \frac{1}{2\pi} \left[ \frac{\sin ((n-m)x)}{n-m} - \frac{\sin ((n+m)x)}{n+m} \right]_{-\pi}^{\pi} \\
= \frac{1}{\pi} \left[ \frac{\sin ((n-m)\pi)}{n-m} - \frac{\sin ((n+m)\pi)}{n+m} \right]
\]

(12.39)

whether \( m \) is equal to or different from \( n \), \( \sin ((n+m)\pi) = 0 \), thus

\[
\langle F_n(x) | F_m(x) \rangle = \frac{1}{\pi} \frac{\sin ((n-m)\pi)}{n-m} = \delta_{nm}.
\]
12.3 Complete Sets of Functions

We recall a set of functions, \( F_n(x) \), form an orthonormal set for all \( x \in (a,b) \), when

\[
\langle F_n(x) | F_m(x) \rangle = \int_a^b F_n^*(x) F_m(x) \, dx = \delta_{nm}. \tag{12.42}
\]

Next we shall study when such kind of set of functions form a complete set for all \( x \in (a,b) \) in the function space. But before we do that, we want to review complete set of vectors in vector space.

**Complete sets of vectors in vector space:** Consider a three dimensional real space \( \mathbb{R}^3 \) in Cartesian coordinate system. We recall the unit vectors \( \hat{x}, \hat{y}, \text{ and } \hat{z} \) or \( \hat{x}_1, \hat{x}_2, \text{ and } \hat{x}_3 \) form an orthonormal set of vectors in 3-D vector space since

\[
\hat{x}_n \cdot \hat{x}_m = \delta_{nm} \tag{12.43}
\]

or using Dirac notation

\[
\langle x_n | x_m \rangle = \delta_{nm}. \tag{12.44}
\]

Any vector \( \vec{r} \in \mathbb{R}^3 \) can be expressed in terms of these three unit vectors

\[
\vec{r} = \sum_{i=1}^{3} r_i \hat{x}_i \tag{12.45}
\]

or using Dirac notation

\[
|\vec{r}\rangle = \sum_{i=1}^{3} r_i |x_i\rangle. \tag{12.46}
\]

Multiplying both sides from the left by \( \langle x_j | \),

\[
\langle x_j | \vec{r} \rangle = \sum_{i=1}^{3} r_i \langle x_j | x_i \rangle. \tag{12.47}
\]

Since the vectors \( |x_i\rangle \) form an orthonormal set of vectors

\[
\langle x_j | \vec{r} \rangle = \sum_{i=1}^{3} r_i \delta_{ji} \Rightarrow r_j = \langle x_j | \vec{r} \rangle. \tag{12.48}
\]

If one of the vectors, \( \hat{x}_i \), is missing, then we can not express any vector \( \vec{r} \in \mathbb{R}^3 \) in terms of the remaining two vectors. But if we have all these three vectors any vector \( \vec{r} \in \mathbb{R}^3 \) can be expressed in terms of these vectors. Then we say that the set of vectors \( \{\hat{x}_i\} = \{\hat{x}_1, \hat{x}_2, \hat{x}_3\} \) is a complete orthonormal set in \( \mathbb{R}^3 \).

We now consider the function space where we have the infinite set of orthonormal functions \( \{F_n(x)\} = \{F_0(x), F_1(x), F_2(x), F_3(x), \ldots\} \) for all \( x \in (a,b) \). These set of functions form a complete set if a function, \( g(x) \), defined
for all $x \in (a, b)$ can be expressed as a linear combination of these orthonormal functions. That means
\begin{equation}
|g(x)\rangle = \sum_{n=0}^{\infty} c_n |F_n(x)\rangle, \tag{12.49}
\end{equation}
where the expansion coefficient, $c_n$, is determined using the property for orthonormal set of functions
\begin{equation}
\langle F_n(x) | F_m(x) \rangle = \int_{a}^{b} F_n^* (x) F_m (x) \, dx = \delta_{nm}. \tag{12.50}
\end{equation}
Multiplying both sides by the bra vector $\langle F_m(x) |$ from the left, we have
\begin{equation}
\langle F_m(x) | g(x) \rangle = \sum_{n=0}^{\infty} c_n \langle F_m(x) | F_n(x) \rangle, \tag{12.51}
\end{equation}
so that using Eq. (12.50), we find
\begin{equation}
\langle F_m(x) | g(x) \rangle = \sum_{n=0}^{\infty} c_n \delta_{nm} = c_m \Rightarrow c_m = \int_{a}^{b} F_m^* (x) g(x) \, dx. \tag{12.52}
\end{equation}

The completeness relation: A set of vectors or functions that form a complete set, using the Dirac equation, are defined by the completeness relation as
\begin{align*}
\sum_{n=0}^{\infty} |x_n\rangle \langle x_n| &= 1, \text{ for a complete set of vectors} \\
\int_{a}^{b} dx |F_n(x)\rangle \langle F_n(x) | &= 1, \text{ for a complete set of functions}
\end{align*}
You will often use completeness relations in quantum mechanics.

**Example 12.3** Any periodic function $g(x)$ defined in the interval $(-\pi, \pi)$ can be expressed in terms of the set of orthonormal function
\begin{equation}
F_n(x) = \frac{1}{\sqrt{\pi}} \sin (nx) = \left\{ \frac{1}{\sqrt{\pi}} \sin (x), \frac{1}{\sqrt{\pi}} \sin (2x), \frac{1}{\sqrt{\pi}} \sin (3x), \ldots \right\} \tag{12.53}
\end{equation}
as
\begin{equation}
|g(x)\rangle = \sum_{n=1}^{\infty} c_n |F_n(x)\rangle \Rightarrow g(x) = \sum_{n=1}^{\infty} \frac{c_n}{\sqrt{\pi}} \sin (nx), \tag{12.54}
\end{equation}
where
\begin{equation}
\Rightarrow c_n = \int_{-\pi}^{\pi} F_n^* (x) g(x) \, dx \Rightarrow c_n = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} \sin (nx) g(x) \, dx. \tag{12.55}
\end{equation}
Thus the set of functions \( \{ F_n(x) \} \) form a complete orthonormal set. We shall see this is nothing but Fourier series expansion. As an example let’s consider the step-function defined by

\[
f(x) = \begin{cases} 
1 & \text{for } 0 < x < \pi \\
-1 & \text{for } -\pi < x < 0 
\end{cases}
\]  

(12.56)

(Example 16.4). The series expansion of this function can be show to be

\[
f(x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin[(2n + 1)x]}{2n + 1}.
\]  

(12.57)

Figure 12.1: Fourier series expansion of the step-function.

This series expansion for different values of terms in the series is shown in Fig. 12.1. It shows the more terms in the series are considered, the more the series expansion be close to the actual function.
12.4 The Legendre Differential Equation

Series substitution method is generally used to solve a linear differential equation of the form

\[ a_n(x) \frac{d^n y}{dx^n} + a_{n-1}(x) \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1(x) \frac{dy}{dx} + a_0(x) = 0. \]  

(12.58)

For the case where the coefficients \( a_n(x) \) are independent of \( x \), we have seen the application of this method by solving the harmonic oscillator problem. Next we shall see how this method is used to determine the solution to the Legendre differential equation where some of the coefficients, \( a_n(x) \), depend on \( x \). The Legendre differential equation, for \( x \in [-1, 1] \), is given by

\[ (1 - x^2) y'' - 2xy' + l(l + 1)y = 0 \]  

(12.59)

where

\[ a_2(x) = 1 - x^2, \quad a_1(x) = -2x, \quad a_0 = l(l + 1) \]

and \( l = 0, 1, 2, 3, \ldots \). The Legendre differential equation is derived from the Laplace’s equation in spherical coordinates. For now we focus on how to find the solution to this differential equation. To this end we may want to rewrite the Legendre differential equation in the form

\[ l(l + 1)y - 2xy' - x^2y'' + y'' = 0. \]  

(12.60)

Suppose the solution to this differential equation, \( y(x) \), exists, one must be able to express the function, \( y(x) \), as a convergent power series given by

\[ y(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots a_n x^n + \cdots \]

\[ \Rightarrow \quad l(l + 1)y = \sum_{n=0}^{\infty} l(l + 1) a_n x^n \]  

(12.60)

The first derivative of this function

\[ y'(x) = \sum_{n=0}^{\infty} a_n \frac{d}{dx} (x^n) = \sum_{n=0}^{\infty} a_n n x^{n-1} \Rightarrow 2xy'(x) = \sum_{n=0}^{\infty} 2n x a_n^n, \]  

(12.61)

The second derivative

\[ y''(x) = \sum_{n=0}^{\infty} a_n n d \frac{d}{dx} (x^{n-1}) = \sum_{n=0}^{\infty} a_n n (n-1) x^{n-2} \]

\[ \Rightarrow \quad x^2 y''(x) = \sum_{n=0}^{\infty} a_n n (n-1) x^n. \]  

(12.62)
Upon examining the first few terms in the series for \( y''(x) \),

\[
y''(x) = \sum_{n=0}^{\infty} a_n n (n-1) x^{n-2} = 0 + 0 + 2a_2 + 3.2a_3x + \ldots \quad (12.63)
\]

we note that

\[
y''(x) = \sum_{m=2}^{\infty} a_m m (m-1) x^{m-2}, \quad (12.64)
\]

where we used a different dummy summation index, \( m \). Now introducing another dummy summation index defined by

\[
n = m - 2 \Rightarrow m = n + 2,
\]

we have

\[
m = 2 \Rightarrow n = m - 2 = 0
\]

so that one can rewrite the second derivative as

\[
y''(x) = \sum_{n=0}^{\infty} (n+2) (n+1) a_{n+2} x^n. \quad (12.65)
\]

Now substituting the expressions for \( y, xy', x^2y'' \), and \( y'' \) into the Legendre differential equation

\[
l(l+1)y - 2xy' - x^2y'' + y'' = 0,
\]

we find

\[
\sum_{n=0}^{\infty} (n+2) (n+1) a_{n+2} x^n + \sum_{n=0}^{\infty} [l(l+1) - n (n-1) - 2n] a_n x^n = 0.
\]

Upon simplifying this expression, we find

\[
\sum_{n=0}^{\infty} \left\{ a_{n+2} (n+2) (n+1) + a_n [l(l+1) - n (n+1)] \right\} x^n = 0. \quad (12.66)
\]

There follows that

\[
a_{n+2} (n+2) (n+1) + a_n [l(l+1) - n (n+1)] = 0
\]

\[
\Rightarrow a_{n+2} = -\frac{l(l+1) - n(n+1)}{(n+2)(n+1)} a_n \quad (12.67)
\]

and noting that

\[
l(l+1) - n(n+1) = l^2 - n^2 + l - n = (l-n)(l+n+1)
\]

one can write

\[
a_{n+2} = -\frac{(l-n)(l+n+1)}{(n+2)(n+1)} a_n. \quad (12.68)
\]
12.4. THE LEGENDRE DIFFERENTIAL EQUATION

This the recursion relation we will next use to determine each \( a_n \), when \( n \) is even or odd.

(a) Even \( n = 2, 4, 6... \)

\[
a_2 = - \frac{(l - 0)(l + 0 + 1)}{(0 + 2)(0 + 1)} a_0 \Rightarrow a_2 = - \frac{l(l + 1)}{2!} a_0 \quad (12.69)
\]

\[
a_4 = - \frac{(l - 2)(l + 2 + 1)}{(2 + 2)(2 + 1)} a_2 \Rightarrow a_4 = \frac{l(l + 1)(l - 2)(l + 3)}{4!} a_0 \quad (12.70)
\]

\[
a_6 = - \frac{(l - 4)(l + 4 + 1)}{(4 + 2)(4 + 1)} a_4 \Rightarrow a_6 = - \frac{l(l + 1)(l - 2)(l + 3)(l - 4)(l + 5)}{6!} a_0 \quad (12.71)
\]

(b) Odd \( n = 1, 3, 5... \)

\[
a_3 = - \frac{(l - 1)(l + 1 + 1)}{(1 + 2)(1 + 1)} a_1 \Rightarrow a_3 = - \frac{l(l + 2)}{3!} a_1 \quad (12.72)
\]

\[
a_5 = - \frac{(l - 3)(l + 3 + 1)}{(3 + 2)(3 + 1)} a_3 \Rightarrow a_5 = \frac{l(l + 1)(l - 2)(l - 3)(l + 4)}{5!} a_1 \quad (12.73)
\]

\[
a_7 = - \frac{(l - 5)(l + 5 + 1)}{(5 + 2)(5 + 1)} a_5 \Rightarrow a_7 = - \frac{l(l + 1)(l - 2)(l + 4)(l - 5)(l + 6)}{7!} a_1. \quad (12.74)
\]

Therefore, the function \( y(x) \) is a sum of two series

\[
y(x) = y_e(x) + y_o(x)
\]

where

\[
y_e(x) = a_0 \left\{ 1 - \frac{l(l + 1)}{2!} x^2 + \frac{l(l + 1)(l - 2)(l + 3)}{4!} x^4 \right. \\
\left. - \frac{l(l + 1)(l - 2)(l + 3)(l - 4)(l + 5)}{6!} x^6 + \ldots \right\} \quad (12.75)
\]
CHAPTER 12. POWER SERIES SOLUTIONS TO DIFFERENTIAL EQUATIONS

and

\[ y_o(x) = a_1 \left\{ x - \frac{(l-1)(l+2)}{3!}x^3 + \frac{(l-1)(l+2)(l-3)(l+4)}{5!}x^5 \right. \]

\[ - \frac{(l-1)(l+2)(l-3)(l+4)(l-5)(l+6)}{7!}x^7 + \ldots \} \]  (12.76)

Note that \( a_0 \) and \( a_1 \) are constants determined from the boundary conditions. Furthermore, we do not know if this series is convergent or divergent. For this function to be a solution to the differential equation it must be a convergent series. The next section focuses on the conditions that must be fulfilled for the series to be a convergent series for \( x \in [-1, 1] \).

12.5 The Legendre Polynomials

The Legendre polynomials are the solution to the Legendre differential equation. These polynomials are determined from, \( y(x) \), that we determined earlier when we impose the condition for convergence to the series. Any function that is a solution to any differential equation must be finite. This means the series form of the function must be convergent in the given domain, \( x \in [-1, 1] \). Using the ratio test

\[ \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} x \right| < 1. \]  (12.77)

and the recursion relation we derived earlier, we have

\[ a_{n+2} = -\frac{(l-n)(l+n+1)}{(n+2)(n+1)} a_n \Rightarrow \lim_{n \to \infty} \left| \frac{a_{n+2}}{a_n} \right| = \lim_{n \to \infty} \left| \frac{(l-n)(l+n+1)}{(n+2)(n+1)} \right| \]  (12.78)

If we determine the interval of convergence for

\[ \lim_{n \to \infty} \left| \frac{a_{n+2}}{a_n} x^2 \right| < 1 \]  (12.79)

it means we have determined the interval of convergence for the even series

\[ y_e(x) = a_0 \left\{ 1 - \frac{l(l+1)}{2!}x^2 + \frac{l(l+1)(l-2)(l+3)}{4!}x^4 \right. \]

\[ - \frac{l(l+1)(l-2)(l+3)(l-4)(l+5)}{6!}x^6 + \ldots \} \]  (12.80)

and the odd series

\[ y_o(x) = a_1 \left\{ x - \frac{(l-1)(l+2)}{3!}x^3 + \frac{(l-1)(l+2)(l-3)(l+4)}{5!}x^5 \right. \]

\[ - \frac{(l-1)(l+2)(l-3)(l+4)(l-5)(l+6)}{7!}x^7 + \ldots \} \].  (12.81)
Therefore, the interval of convergence for the function
\[ y(x) = y_e(x) + y_o(x) \] (12.82)
would be the intervals of convergence that is common to both series. Since each of the coefficients in both the even and odd series satisfy the condition
\[
\lim_{n \to \infty} \left| \frac{a_{n+2}}{a_n} x^2 \right| = \lim_{n \to \infty} \left| \frac{(l-n)(l+n+1)}{(n+2)(n+1)} x^2 \right| < 1
\]
\[
\Rightarrow \lim_{n \to \infty} \left| \frac{(l-n)(l+n+1)}{(n+2)(n+1)} x^2 \right| = \lim_{n \to \infty} |x| < 1 \] (12.83)
we can conclude that the series
\[ y(x) = y_e(x) + y_o(x) \] (12.84)
is convergent in the interval \(-1 < x < 1\). However, the ratio test fails at \(x = \pm 1\), we need to examine the series at the boundaries \(x = \pm 1\). For \(x = \pm 1\), we find for the even series
\[
y_e(x) = a_0 \left\{ 1 - \frac{l(l+1)}{2!} + \frac{l(l+1)(l-2)(l+3)}{4!} \right.
\]
\[
- \frac{l(l+1)(l-2)(l+3)(l-4)(l+5)}{6!} + \ldots \} \] (12.85)
and the odd series
\[
y_o(x) = a_1 \left\{ 1 - \frac{(l-1)(l+2)}{3!} + \frac{(l-1)(l+2)(l-3)(l+4)}{5!} \right.
\]
\[
- \frac{(l-1)(l+2)(l-3)(l+4)(l-5)(l+6)}{7!} + \ldots \} \] (12.86)
where we included the factor \(\pm 1\) in the odd series in the constant \(a_1\). In the Legendre differential equation, we recall that, \(l = 0, 1, 2, 3, \ldots\) so that if we set these values one finds for \(l = 0\),
\[ y_e(x) = a_0 \] (12.87)
and
\[ y_o(x) = a_1 \left\{ 1 + \frac{1 \times 2}{3!} + \frac{1 \times 2 \times 3 \times 4}{5!} + \frac{1 \times 2 \times 3 \times 4 \times 5 \times 6}{7!} + \ldots \} \]
\[ = a_1 \left\{ 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \ldots + \frac{1}{2n+1} + \ldots \right\} = a_1 \sum_{n=0}^{\infty} \frac{1}{2n+1} \] (12.88)
Is this series a convergent series? Using the integral test, we find
\[
\int_{0}^{\infty} \frac{dn}{2n+1} = \frac{1}{2} \ln (2n+1) \bigg|_{0}^{\infty} = \infty, \] (12.89)
which means the series is a divergent series. Since we are looking for a well defined function for which the series is convergent for all \( x \) in the interval, \(-1 \leq x \leq 1\), we must set, \( a_1 = 0 \), so that

\[
y_o(x) = a_1 \sum_{n=0}^{\infty} \frac{1}{2n + 1} = 0
\]

(12.90)

and the solution to the Legendre differential equation, for \( l = 0 \) is

\[
y(x) = y_e(x) = a_0
\]

(12.91)

For \( l = 1 \), at \( x = \pm 1 \), we have

\[
y_e(\pm 1) = a_0 \left\{ 1 - \frac{1(1+1)}{2!} + \frac{1(1+1)(1-2)(1+3)}{4!} - \frac{1(1+1)(1-2)(1+3)(1-4)(1+5)}{6!} + \ldots \right\}
\]

\[
= a_0 \left\{ 1 - \frac{2}{2!} - \frac{1(2)(4)}{4!} - \frac{1(2)(4)(3)(6)}{6!} + \ldots \right\}
\]

\[
= a_0 \left\{ \frac{1}{3} + \frac{1}{5} + \frac{1}{7} - \frac{1}{2n+3} + \ldots \right\}
\]

\[
\Rightarrow y_e(\pm 1) = a_0 \sum_{n=0}^{\infty} \frac{1}{2n + 3}
\]

(12.92)

which is a divergent series and we must have, \( a_0 = 0 \), so that

\[
y_e(\pm 1) = 0.
\]

(12.93)

On the other hand, for the odd series for all \( x \) (i.e. \(-1 \leq x \leq 1\)) at \( l = 1 \), we find

\[
y_o(x) = a_1 \left\{ 1 - \frac{(1-1)(1+2)}{3!} + \frac{(1-1)(1+2)(1-3)(1+4)}{5!} - \frac{(1-1)(1+2)(1-3)(1+4)(1-5)(1+6)}{7!} + \ldots \right\}
\]

\[
y_o(x) = a_1 x,
\]

(12.94)

which leads to

\[
y(x) = y_o(x) = a_1 x \text{ for } l = 1.
\]

(12.95)

For the same reason discussed above it can be shown that the solution to the Legendre differential equation must then be given by

\[
y(x) = \begin{cases} y_e(x), & l = \text{even} \\ y_o(x), & l = \text{odd} \end{cases}
\]

(12.96)
12.6. THE GENERATING FUNCTION FOR THE LEGENDRE POLYNOMIALS

where

\[
y_e(x) = a_0 \left\{ 1 - \frac{l(l+1)}{2!} x^2 + \frac{l(l+1)(l-2)(l+3)}{4!} x^4 - \frac{l(l+1)(l-2)(l+3)(l-4)(l+5)}{6!} x^6 + \ldots \right\},
\]

\[
y_o(x) = a_1 \left\{ x - \frac{(l-1)(l+2)}{3!} x^3 + \frac{(l-1)(l+2)(l-3)(l+4)}{5!} x^5 - \frac{(l-1)(l+2)(l-3)(l-5)(l+6)}{7!} x^7 + \ldots \right\}.
\]

For example, \( l = 2 \) and \( l = 3 \),

\[ l = 2 \Rightarrow y(x) = a_0 \left( 1 - 3x^2 \right), \quad l = 3 \Rightarrow y(x) = a_1 \left( x - \frac{5}{3} x^3 \right) \]

If we chose \( a_0 \) and \( a_1 \) such that \( y(x) = 1 \) at \( x = 1 \), we find

\[
\begin{align*}
  l &= 0 \Rightarrow y_0(x) = P_0(x) = 1, \\
  l &= 1 \Rightarrow y_1(x) = P_1(x) = x, \\
  l &= 2 \Rightarrow y_2(x) = P_2(x) = \frac{1}{2} (3x^2 - 1), \\
  l &= 3 \Rightarrow y_3(x) = P_3(x) = \frac{1}{2} (5x^3 - 3x).
\end{align*}
\]

These polynomials, \( P_n(x) \), are called the Legendre Polynomials.

12.6 The Generating Function for the Legendre Polynomials

We recall that the Legendre polynomials for \( l = 0, 1, 2, 3 \) are

\[
P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2} (3x^2 - 1), \quad P_3(x) = \frac{1}{2} (5x^3 - 3x).
\]

In this section we are interested in finding a function that generates these Polynomials. To this end, we shall consider the function, which is referred as the Generating function,

\[
\phi(x, h) = (1 - 2xh + h^2)^{-1/2}.
\]

We want to examine the Taylor series expansion for the function

\[
\phi(x, h) = \left( 1 - 2xh + h^2 \right)^{-1/2}
\]

(12.103)
CHAPTER 12. POWER SERIES SOLUTIONS TO DIFFERENTIAL EQUATIONS

about, \( h = 0 \) which is given by

\[
\phi (x, h) = \sum_{l=0}^{\infty} \frac{1}{l!} \frac{d^l}{dh^l} \phi (x, h) \bigg|_{h=0} h^l = \phi (x, 0) + \left[ \frac{d}{dh} \phi (x, 0) \right] \frac{h}{1!} + \left[ \frac{d^2}{dh^2} \phi (x, 0) \right] \frac{h^2}{2!} + \frac{d^3}{dh^3} \phi (x, 0) \right] \frac{h^3}{3!} + \cdots
\]  

(12.104)

For \( l = 0 \)

\[
\phi (x, 0) = 1 = P_0 (x),
\]

(12.105)

\[
l = 1
\]

\[
\frac{d}{dh} \phi (x, h) = \frac{x - h}{(1 - 2xh + h^2)^{3/2}} \Rightarrow \frac{d}{dh} \phi (x, h) \bigg|_{h=0} = x = P_1 (x),
\]

(12.106)

\[
l = 2
\]

\[
\frac{d^2}{dh^2} \phi (x, h) = \frac{d}{dh} \left[ \frac{x - h}{(1 - 2xh + h^2)^{3/2}} \right]
\]

\[= - \frac{1}{(1 - 2xh + h^2)^{3/2}} \frac{3 (x - h)^2}{(1 - 2xh + h^2)^{5/2}}
\]

\[
\frac{d^2}{dh^2} \phi (x, h) \bigg|_{h=0} = -1 + 3x^2 = 2! \left( \frac{1}{2} (3x^2 - 1) = 2! P_2 (x),
\]

(12.107)

\[
l = 3
\]

\[
\frac{d^3}{dh^3} \phi (x, h) = \frac{d}{dh} \left[ - \frac{1}{(1 - 2xh + h^2)^{3/2}} + \frac{3 (x - h)^2}{(1 - 2xh + h^2)^{5/2}} \right]
\]

\[= -3 (x - h) \frac{3 \times 2 (x - h)}{(1 - 2xh + h^2)^{5/2}} + \frac{3 \times 5 (x - h)^3}{(1 - 2xh + h^2)^{7/2}}
\]

\[
\frac{d^3}{dh^3} \phi (x, h) \bigg|_{h=0} = 15x^3 - 9x = 3! \frac{1}{2} (5x^3 - 3x) = 3! P_3 (x),
\]

(12.108)

we can then conclude for the \( l^{th} \) term in the series

\[
\frac{d^l}{dh^l} \phi (x, h) \bigg|_{h=0} = l! P_l (x)
\]

(12.109)

Therefore, the Taylor series expansion for the function

\[
\phi (x, h) = (1 - 2xh + h^2)^{-1/2}
\]

(12.110)

about \( h = 0 \) can be written as

\[
\phi (x, h) = \sum_{l=0}^{\infty} \frac{1}{l!} \frac{d^l}{dh^l} \phi (x, h) \bigg|_{h=0} h^l = \sum_{l=0}^{\infty} P_l (x) h^l.
\]

(12.111)

The function \( \phi (x, h) \) is the generating function for the Legendre polynomials.
Example 12.4 Consider a point charge $Q$ located at the position $\vec{r}'$ (the source point). Find an expression for the electrostatic potential $V(r)$ at the point $P$ located at the position $\vec{r}$ (the field point). (Fig. 12.6)

Solution: For a point charge located at the position described by the vector $\vec{r}'$, the electric potential at the point at position $\vec{r}$ is inversely proportional to the distance between the charge position and the point $P$ (|$\vec{r} - \vec{r}'$|) and directly proportional to the charge. It can be expressed as

$$V(r) = \frac{Q}{4\pi\varepsilon_0} \frac{1}{|r - r'|}.$$  

(12.112)

Noting that

$$\frac{1}{|\vec{r} - \vec{r}'|} = (r^2 + r'^2 - 2rr' \cos \theta)^{-1/2}$$  

(12.113)

For $r' < r$ we may write

$$\frac{1}{|\vec{r} - \vec{r}'|} = \frac{1}{r} \left[ 1 + \left( \frac{r'}{r} \right)^2 - 2 \frac{r'}{r} \cos \theta \right]^{-1/2}.$$  

(12.114)

If we introduce the variables

$$h = \frac{r'}{r}, \quad x = \cos \theta$$  

(12.115)

we have

$$\frac{1}{|\vec{r} - \vec{r}'|} = \frac{1}{r} \left( 1 - 2xh + h^2 \right)^{1/2}$$  

(12.116)

so that using

$$(1 - 2xh + h^2)^{-1/2} = \sum_{l=0}^{\infty} P_l(x) h^l$$  

(12.117)
for
\[ h = \frac{r'}{r}, \quad x = \cos \theta \] (12.118)

the expression for the electric potential becomes
\[ V(r) = \frac{Q}{4\pi \varepsilon_0 \left| r - r' \right|} = \frac{Q}{4\pi \varepsilon_0 r'} \sum_{l=0}^{\infty} P_l (\cos \theta) \left( \frac{r'}{r} \right)^l \] (12.119)

when \( r' < r \). On the other hand when \( r' > r \), we must write
\[ \frac{1}{\left| r - r' \right|} = \frac{1}{r'} \left[ 1 + \left( \frac{r}{r'} \right)^2 - 2 \frac{r}{r'} \cos \theta \right]^{-1/2}. \] (12.120)

or
\[ \frac{1}{\left| r - r' \right|} = \frac{1}{r'} \left( 1 - 2xh + h^2 \right)^{1/2} \] (12.121)

where
\[ h = \frac{r'}{r}, \quad x = \cos \theta. \] (12.122)

In this case the potential is given by
\[ V(r) = \frac{Q}{4\pi \varepsilon_0 \left| r - r' \right|} = \frac{Q}{4\pi \varepsilon_0 r'} \sum_{l=0}^{\infty} P_l (\cos \theta) \left( \frac{r'}{r} \right)^l \] (12.123)

The Rodrigues’ Formula: it gives Legendre Polynomials:
\[ P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} \left( x^2 - 1 \right)^l \] (12.124)

**Example 12.5** Use Rodrigues’ formula to find the equation for \( P_2(x) \).

**Solution:** Using Rodrigues’ Formula \( P_2(x) \) can be expressed as
\[ P_2(x) = \frac{1}{2^2 2!} \frac{d^2}{dx^2} \left( x^2 - 1 \right)^2 \] (12.125)

which can be simplified as follows:
\[ P_2(x) = \frac{1}{2^2 2!} \frac{d^2}{dx^2} \left( x^4 - 2x^2 + 1 \right) = \frac{1}{2^2 2!} \frac{d}{dx} \left( 4x^3 - 4x \right) \]
\[ \Rightarrow P_2(x) = \frac{1}{2^3} \left[ 12x^2 - 4 \right] = \frac{4}{2^3} \left[ 3x^2 - 1 \right] = \frac{1}{2} \left[ 3x^2 - 1 \right] \] (12.126)

**Leibniz’ Rule:** it is useful when we need to find a higher order derivatives of a product (for example when we apply Rodrigues’ Formula). It is given by
\[ \frac{d^N}{dx^N} (uv) = \sum_{n=0}^{N} \binom{N}{n} \left( \frac{d^n u}{dx^n} \right) \left( \frac{d^{N-n} v}{dx^{N-n}} \right) \] (12.127)
Example 12.6 Use Leibniz’ Rule to find

\[
\frac{d^8}{dx^8} (x^2 e^{2x}).
\]  

(12.128)

Solution: For \( N = 8 \) Leibniz’ Rule can be expressed as

\[
\frac{d^8}{dx^8} (uv) = \sum_{n=0}^{8} \binom{8}{n} \left( \frac{d^n u}{dx^n} \frac{d^{8-n} v}{dx^{8-n}} \right).
\]  

(12.129)

For \( u = x^2 \), we have

\[
\frac{d^0 u}{dx^0} = x^2, \quad \frac{d^1 u}{dx^1} = 2x, \quad \frac{d^2 u}{dx^2} = 2, \quad \frac{d^n u}{dx^n} = 0 \text{ for } n \geq 3
\]  

(12.130)

so that

\[
\frac{d^8}{dx^8} (uv) = \sum_{n=0}^{8} \binom{8}{n} \left( \frac{d^n u}{dx^n} \frac{d^{8-n} v}{dx^{8-n}} \right) + \binom{8}{1} \frac{d^1 u}{dx^1} \frac{d^7 v}{dx^7} + \binom{8}{2} \frac{d^2 u}{dx^2} \frac{d^6 v}{dx^6}
\]  

(12.131)

Noting that for \( v = e^{2x} \)

\[
\frac{d^n v}{dx^n} = 2^n e^{2x}
\]  

(12.132)

we find

\[
\frac{d^8}{dx^8} (uv) = x^2 (2^8 e^{2x}) + 16x (2^7 e^{2x}) + 56 (2^6 e^{2x}) = 2^8 (x^2 + 8x + 14) e^{2x}
\]  

\[= \frac{d^8}{dx^8} (uv) = (256x^2 + 2048x + 3584) e^{2x}
\]  

(12.133)

A Recursion Relation: We recall the generating function

\[
(1 - 2xh + h^2)^{-1/2} = \sum_{l=0}^{\infty} P_l (x) h^l
\]  

(12.134)

so that

\[
\frac{d}{dh} (1 - 2xh + h^2)^{-1/2} = \frac{d}{dh} \sum_{l=0}^{\infty} P_l (x) h^l
\]

\[
\Rightarrow \frac{(x - h)}{(1 - 2xh + h^2)^{1/2}} = \sum_{l=0}^{\infty} lP_l (x) h^{l-1}.
\]  

(12.135)
If we use
\[
\frac{1}{\sqrt{1 - 2xh + h^2}} = \sum_{l=0}^{\infty} P_l(x) h^l
\]
we can write the above expression as
\[
\frac{(x - h) \sum_{l=0}^{\infty} P_l(x) h^l}{1 - 2xh + h^2} = \sum_{l=0}^{\infty} lP_l(x) h^{l-1}
\]
\[\Rightarrow (x - h) \sum_{l=0}^{\infty} P_l(x) h^l = (1 - 2xh + h^2) \sum_{l=0}^{\infty} lP_l(x) h^{l-1} \]
\[\Rightarrow \sum_{l=0}^{\infty} P_l(x) h^l - \sum_{l=0}^{\infty} P_l(x) h^{l+1} = \sum_{l=0}^{\infty} lP_l(x) h^{l-1} - 2x \sum_{l=0}^{\infty} lP_l(x) h^l + \sum_{l=0}^{\infty} lP_l(x) h^{l+1} \]
\[\Rightarrow \sum_{l=0}^{\infty} lP_l(x) h^{l-1} - \sum_{l=0}^{\infty} x (1 + 2l) P_l(x) h^l + \sum_{l=0}^{\infty} (1 + l) P_l(x) h^{l+1} = 0 \]
(12.137)

Noting that
\[
\sum_{l=0}^{\infty} lP_l(x) h^{l-1} = 1P_1(x) + 2P_2(x) h^1 + 2P_3(x) h^2 + 4P_4(x) h^3 + \ldots 
\[
= \sum_{l=0}^{\infty} (l + 1) P_{l+1}(x) h^l
\]
(12.138)

and
\[
\sum_{l=0}^{\infty} (1 + l) P_l(x) h^{l+1} = P_0(x) h + 2P_1(x) h^2 + 3P_2(x) h^3 + 4P_3(x) h^4 + 5P_4(x) h^5 + \ldots = 0 \times P_{-1}(x) + P_0(x) h + 2P_1(x) h^2 + 3P_2(x) h^3 + P_3(x) h^4 + \sum_{l=0}^{\infty} (1 + l) P_l(x) h^{l+1}
\]
\[\Rightarrow \sum_{l=0}^{\infty} (1 + l) P_l(x) h^{l+1} = \sum_{l=0}^{\infty} lP_{l-1}(x) h^l, \]
(12.139)

we can write
\[
\sum_{l=0}^{\infty} (l + 1) P_{l+1}(x) h^l - \sum_{l=0}^{\infty} x (1 + 2l) P_l(x) h^l + \sum_{l=0}^{\infty} lP_{l-1}(x) h^l = 0
\]
\[\Rightarrow \sum_{l=0}^{\infty} [(l + 1) P_{l+1}(x) - x (1 + 2l) P_l(x) + lP_{l-1}(x)] h^l = 0 \]
(12.140)
12.7. LEGENDRE SERIES

There follows that

\[(l + 1) P_{l+1}(x) - x(1 + 2l) P_l(x) + lP_{l-1}(x) = 0 \]
\[\Rightarrow (l + 1) P_{l+1}(x) = x(1 + 2l) P_l(x) - lP_{l-1}(x). \quad (12.141)\]

This is the recursion relation for the Legendre polynomials.

### 12.7 Legendre Series

**Orthogonality of the Legendre polynomials:** Legendre polynomials form an orthogonal set of functions. The orthogonality conditions is given by

\[\langle P_l(x)|P_m(x)\rangle = \frac{2}{2l+1} \delta_{lm} \quad (12.142)\]

for \(x \in [-1, 1]\).

**Completeness of the Legendre Polynomials:** The Legendre polynomials also form complete set on the interval \([-1, 1]\). That means any function \(f(x)\) can be expressed as a linear combination of the Legendre polynomials

\[f(x) = \sum_{m=0}^{\infty} a_m |P_m(x)|. \quad (12.143)\]

The expression for the expansion coefficients is determined using the orthogonality property of the Legendre polynomials. Multiplying both sides from the left by \(\langle P_l(x)\rangle\), we can write

\[\langle P_l(x)|f(x)\rangle = \sum_{m=0}^{\infty} a_m \langle P_l(x)|P_m(x)\rangle = \sum_{m=0}^{\infty} a_m \frac{2}{2l+1} \delta_{lm}\]
\[\Rightarrow \langle P_l(x)|f(x)\rangle = a_l \frac{2}{2l+1} \Rightarrow a_l = \frac{2l+1}{2} \langle P_l(x)|f(x)\rangle. \quad (12.144)\]

Recalling that

\[\langle P_l(x)|f(x)\rangle = \int_{-1}^{1} P_l^*(x) f(x) dx \quad (12.145)\]

and \(P_l(x)\) is real, the expression for the expansion coefficients is expressible as

\[a_l = \frac{2l+1}{2} \int_{-1}^{1} P_l(x) f(x) dx. \quad (12.146)\]

**Example 12.7** Expand the function \(f(x)\) in a Legendre series, for

\[f(x) = \begin{cases} 
-1 & -1 \leq x \leq 0 \\
+1 & 0 \leq x \leq 1
\end{cases} \quad (12.147)\]
Solution: The Legendre series is given by

\[
    f(x) = \sum_{m=0}^{\infty} a_m P_m(x)
\]  \hspace{1cm} (12.148)

where

\[
    a_l = \frac{2l+1}{2} \int_{-1}^{1} P^*_l(x) f(x) \, dx.
\]  \hspace{1cm} (12.149)

The function has different values in the interval \((-1, 0)\) and \((0, 1)\). Hence
the expansion coefficients, \(a_l\) can be expressed as

\[
a_l = \frac{2l + 1}{2} \left[ \int_{-1}^{0} P_l(x) f(x) \, dx + \int_{0}^{1} P_l(x) f(x) \, dx \right]
\]

\[
= \frac{2l + 1}{2} \left[ \int_{0}^{1} P_l(x) \, dx - \int_{-1}^{0} P_l(-x) \, dx \right]
\]

\[
= \frac{2l + 1}{2} \left[ \int_{0}^{1} P_l(-x) \, dx + \int_{0}^{1} P_l(x) \, dx \right]
\]

\[
= \frac{2l + 1}{2} \left[ - \int_{0}^{1} P_l(-x) \, dx + \int_{0}^{1} P_l(x) \, dx \right]
\]

\[
\Rightarrow a_l = \frac{2l + 1}{2} \left[ \int_{0}^{1} [P_l(x) - P_l(-x)] \, dx \right]. \quad (12.150)
\]

Using the Rodrigues formula for the Legendre polynomials

\[
P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} \left( x^2 - 1 \right)^l \quad (12.151)
\]

if we replace \(x\) by \(-x\), we have

\[
P_l(-x) = \frac{1}{2^l l!} \frac{d^l}{d(-x)^l} \left( x^2 - 1 \right)^l = (-1)^l \frac{1}{2^l l!} \frac{d^l}{dx^l} \left( x^2 - 1 \right)^l \quad (12.152)
\]

we leads to

\[
P_l(-x) = \begin{cases} 
P_l(x) & l = \text{even} \\
-P_l(x) & l = \text{odd}
\end{cases} \quad (12.153)
\]

Applying this result we find for \(a_l\)

\[
a_l = \begin{cases} 
(2l + 1) \int_{0}^{1} P_l(x) \, dx, & l = \text{odd} \\
0, & l = \text{even}
\end{cases} \quad (12.154)
\]

Using

\[
P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{1}{2} (3x^2 - 1), P_3(x) = \frac{1}{2} (5x^3 - 3x).
\]
we find
\[ a_1 = 3 \int_0^1 x \, dx = \frac{3}{2}, \quad a_3 = \frac{7}{2} \int_0^1 (5x^3 - 3x) \, dx = -\frac{7}{3} \tag{12.156} \]

Therefore, the Legendre series is
\[ f(x) = \frac{3}{2} P_1(x) - \frac{7}{3} P_2(x) + \ldots \tag{12.157} \]

**Least-Squares Fit and Legendre Series:** Suppose we are given a curve or a graph that is described by a function \( f(x) \) in the interval \((-1,1)\) and we want to find a polynomial
\[ g_n(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 = \sum_{k=0}^n a_k x^k \tag{12.158} \]

that can best fit to the curve described by the function \( f(x) \). Then the polynomial \( g_n(x) \) can best fit the curve described by \( f(x) \) when the integral
\[ \int_{-1}^1 [f(x) - g_n(x)]^2 \, dx \tag{12.159} \]

a minimum. This polynomial is nothing but the \( n^{th} \) order Legendre polynomial.
12.8 The Associated Legendre Differential Equation

The Associated Legendre Differential Equation is given by

\[
\frac{1}{x^2} y'' - \frac{2y'}{x} + \left[ l(l+1) - \frac{m^2}{1-x^2} \right] y = 0,
\]

where \(m^2 \leq l^2\). For \(m = 0\), we find the Legendre Differential equation

\[
(1 - x^2) y'' - 2xy' + l(l+1) y = 0
\]

the solution of which we found to be the Legendre polynomials

\[
y(x) = P_l(x).
\]

We can solve this DE using series substituting method. However, here we use a different approach. To this end we introduce a transformation of variable defined by

\[
y(x) = (1 - x^2)^{m/2} u(x) .
\]

(12.160)

Upon differentiating this with respect to \(x\)

\[
y'(x) = (1 - x^2)^{m/2} u'(x) - m (1 - x^2)^{(m-2)/2} xu(x)
\]

(12.161)

and

\[
y''(x) = (1 - x^2)^{m/2} u''(x) - m (1 - x^2)^{(m-2)/2} xu'(x)
\]

\[-m (1 - x^2)^{(m-2)/2} u(x) + m (m-2) (1 - x^2)^{(m-4)/2} x^2 u(x)
\]

\[-m (1 - x^2)^{(m-2)/2} xu'(x)
\]

\[\Rightarrow y''(x) = (1 - x^2)^{m/2} u''(x) - 2m (1 - x^2)^{(m-2)/2} xu'(x)
\]

\[-m (1 - x^2)^{(m-2)/2} u(x) + m (m-2) (1 - x^2)^{(m-4)/2} x^2 u(x)
\]

(12.162)

There follows that

\[
(1 - x^2) y''(x) = (1 - x^2)^{(m+2)/2} u''(x) - 2m (1 - x^2)^{m/2} xu'(x)
\]

\[-m (1 - x^2)^{m/2} u(x) + m (m-2) (1 - x^2)^{(m-2)/2} x^2 u(x).
\]

(12.163)

We recall

\[y(x) = (1 - x^2)^{m/2} xu(x)
\]

(12.164)

so that

\[
\left[ l(l+1) - \frac{m^2}{1-x^2} \right] y = l(l+1) (1 - x^2)^{m/2} u(x) - m^2 (1 - x^2)^{-m/2} u(x).
\]

(12.165)
and using
\[ y'(x) = \left(1 - x^2\right)^{m/2} u'(x) - m \left(1 - x^2\right)^{(m-2)/2} xu(x) \]  
(12.166)
we have
\[ 2xy'(x) = 2 \left(1 - x^2\right)^{m/2} xu'(x) - 2m \left(1 - x^2\right)^{(m-2)/2} x^2 u(x). \]  
(12.167)
Using the results above we can write the DE
\[ (1 - x^2) y'' - 2xy' + \left[l(l+1) - \frac{m^2}{1-x^2}\right] y = 0 \]  
(12.168)
as
\[
\begin{align*}
(1 - x^2)^{(m+2)/2} u''(x) & - 2m(1 - x^2)^{m/2} xu'(x) - m(1 - x^2)^{m/2} u(x) \\
+ m(m-2) (1 - x^2)^{(m-2)/2} x^2 u(x) - 2(1 - x^2)^{m/2} xu'(x) & \\
+ 2m(1 - x^2)^{(m-2)/2} x^2 u(x) + l(l+1) (1 - x^2)^{m/2} u(x) & \\
- m^2 (1 - x^2)^{-m/2} u(x) & = 0.
\end{align*}
\]  
(12.169)
Dividing the entire equation by \((1 - x^2)^{m/2}\), we find
\[
\begin{align*}
(1 - x^2) u''(x) & - 2mxu'(x) - mu(x) + m(m-2) \frac{x^2}{1-x^2} u(x) \\
- 2xu'(x) & + 2m \frac{x^2}{1-x^2} u(x) + l(l+1) u(x) - m^2 \frac{1}{1-x^2} u(x) = 0 \\
\Rightarrow (1 - x^2) u''(x) & - 2mxu'(x) - 2xu'(x) + l(l+1) u(x) - mu(x) \\
+ m(m-2) \frac{x^2}{1-x^2} u(x) & + \frac{2mx^2}{1-x^2} u(x) - \frac{m^2}{1-x^2} u(x) = 0
\end{align*}
\]  
(12.170)
and the associated differential equation can be put in the form
\[
(1 - x^2) u''(x) - 2 (m+1) xu'(x) + \left[l(l+1) - m(m+1)\right] u(x) = 0. \]  
(12.172)
To find the general solution for the values of \(m\), we shall examine the differential equation for the first three values of \(m\).
12.8. THE ASSOCIATED LEGENDRE DIFFERENTIAL EQUATION

(a) \(m=0\): we may rewrite the differential equation

\[
(1 - x^2) u''(x) - 2xu'(x) + l(l + 1)u(x) = 0. \tag{12.173}
\]

in the form

\[
(1 - x^2) \frac{d^2f_0(x)}{dx^2} - 2xf_0(x) + l(l + 1)f_0(x) = 0. \tag{12.174}
\]

which is the Legendre differential equation and the solution is given by the Legendre polynomials

\[u(x) = f_0(x) = P_l(x) = \frac{d^l P_l(x)}{dx^l}. \tag{12.175}\]

(b) \(m=1\): the differential equation

\[
(1 - x^2) u''(x) - 2xu'(x) + [l(l + 1) - m(m + 1)]u(x) = 0. \tag{12.176}
\]

becomes

\[
(1 - x^2) u''(x) - 4xu'(x) + [l(l + 1) - 2]u(x) = 0. \tag{12.177}
\]

If we differentiate the differential equation we found for \(m = 0\) with respect to \(x\), we find

\[
\frac{d}{dx} \left\{ (1 - x^2) \frac{d^2f_0(x)}{dx^2} - 2xf_0(x) + l(l + 1)f_0(x) = 0 \right\}
\]

\[
\Rightarrow (1 - x^2) \frac{d^3f_0(x)}{dx^3} - 2x \frac{d^2f_0(x)}{dx^2} - 2x \frac{d^2f_0(x)}{dx^2} - 2 \frac{df_0(x)}{dx} + l(l + 1) \frac{df_0(x)}{dx} = 0
\]

\[
\Rightarrow (1 - x^2) \frac{d^3f_0(x)}{dx^3} - 4x \frac{d^2f_0(x)}{dx^2} + (l(l + 1) - 2) \frac{df_0(x)}{dx} = 0 \tag{12.178}
\]

Introducing the transformation defined by

\[f_1(x) = \frac{df_0(x)}{dx} = \frac{d^l P_l(x)}{dx^l}\]

we find

\[
(1 - x^2) \frac{d^2f_1(x)}{dx^2} - 4x \frac{df_1(x)}{dx} + (l(l + 1) - 2) f_1(x) = 0
\]

and comparing this with the differential equation for \(m=1\)

\[
(1 - x^2) \frac{d^2u(x)}{dx^2} - 4x \frac{du(x)}{dx} + [l(l + 1) - 2] u(x) = 0. \tag{12.179}
\]

one can easily see that

\[u(x) = f_1(x) = \frac{d^l P_l(x)}{dx^l}.\]
CHAPTER 12. POWER SERIES SOLUTIONS TO DIFFERENTIAL EQUATIONS

(c) \( m=2 \): the differential equation

\[
(1 - x^2) u''(x) - 2 (m + 1) x u'(x) + [l (l + 1) - m (m + 1)] u(x) = 0.
\]

(12.180)

becomes

\[
(1 - x^2) u''(x) - 6 x u'(x) + [l (l + 1) - 6] u(x) = 0.
\]

(12.181)

Now let’s differentiate the differential equation we found for \( m = 1 \) with respect to \( x \)

\[
\frac{d}{dx} \left\{ (1 - x^2) \frac{d^2 f_1(x)}{dx^2} - 4x \frac{df_1(x)}{dx} + (l (l + 1) - 2) f_1(x) = 0 \right\}
\]

\[
\Rightarrow (1 - x^2) \frac{d^3 f_1(x)}{dx^3} - 6x \frac{d^2 f_1(x)}{dx^2} + (l (l + 1) u'(x) - 6) \frac{df_1(x)}{dx} = 0
\]

Introducing the transformation defined by

\[
f_2(x) = \frac{df_1(x)}{dx} = \frac{d^2 P_l(x)}{dx^2}
\]

we find

\[
(1 - x^2) \frac{d^2 f_2(x)}{dx^2} - 6x \frac{df_2(x)}{dx} + [l (l + 1) - 6] f_2(x) = 0.
\]

(12.182)

which is the same as the differential equation we found for \( m = 2 \)

\[
(1 - x^2) \frac{d^2 u(x)}{dx^2} - 6x \frac{du(x)}{dx} + [l (l + 1) - 6] u(x) = 0.
\]

(12.183)

and the solution is given by

\[
u(x) = f_2(x) = \frac{df_1(x)}{dx} = \frac{d^2 P_l(x)}{dx^2}
\]

Following the same procedure one can easily establish that the solution to the differential equation

\[
(1 - x^2) u''(x) - 2 (m + 1) x u'(x) + [l (l + 1) - m (m + 1)] u(x) = 0.
\]

can be expressed as

\[
u(x) = \frac{d^m P_l(x)}{dx^m}.
\]

(12.184)

Substituting this into

\[
y(x) = (1 - x^2)^{m/2} \ u(x)
\]

(12.185)

the solution to the Associated Legendre Differential equation

\[
(1 - x^2) y'' - 2 x y' + \left[ l (l + 1) - \frac{m^2}{1 - x^2} \right] y = 0.
\]

(12.186)

can then be expressed as

\[
y(x) = (1 - x^2)^{m/2} \frac{d^m P_l(x)}{dx^m}
\]

(12.187)
12.9. SPHERICAL HARMONICS AND THE ADDITION THEOREM

The Associated Legendre Functions: The solutions to the Associated Legendre Differential equation are the Associated Legendre functions and are given by

\[ P^m_l (x) = \left( 1 - x^2 \right)^{m/2} \frac{d^m}{dx^m} P_l (x). \] (12.188)

Note that when \( m = 0 \), we find the Legendre Polynomials

\[ P^0_l (x) = P_l (x). \] (12.189)

Orthogonality of the Associated Legendre Functions:

\[ \langle P^m_l | P^m_l \rangle = \int_{-1}^{1} P^m_l (x) P^m_l (x) \, dx = \frac{2}{2l + 1} \frac{(l + m)!}{(l - m)!} \delta_{lm}, \] (12.190)

where we used that fact that the associated Legendre polynomials are real functions.

12.9 Spherical Harmonics and the addition theorem

Consider the second order differential equation

\[ d^2 \Phi (\varphi) + m^2 \Phi (\varphi) = 0, \] (12.191)

the solution which is given by

\[ \Phi (\varphi) = A (m) e^{im\varphi}. \] (12.192)

Let’s also consider the associated Legendre differential equation

\[ (1 - x^2) \frac{d^2 y (x)}{dx^2} - 2x \frac{dy (x)}{dx} + \left[ l (l + 1) - \frac{m^2}{1 - x^2} \right] y (x) = 0 \] (12.193)

which we determined the solution to be

\[ y (x) = P^m_l (x) = \left( 1 - x^2 \right)^{m/2} \frac{d^m}{dx^m} P_l (x). \] (12.194)

Now introducing the transformation defined by

\[ y (x) = \Theta (\theta), x = \cos (\theta) \Rightarrow \sin^2 (\theta) = 1 - x^2, dx = -\sin (\theta) d\theta \] (12.195)
we have
\[
\frac{dy(x)}{dx} = \frac{d\Theta}{d\theta} \frac{d\theta}{dx} = -\frac{1}{\sin(\theta)} \frac{d\Theta}{d\theta}
\] (12.196)
so that
\[
\frac{d^2y(x)}{dx^2} = -\frac{d\Theta}{d\theta} \left[ \frac{1}{\sin(\theta)} \left( \frac{d\theta}{dx} \right) \right] - \frac{1}{\sin(\theta)} \frac{d^2\Theta}{d\theta^2} \left( \frac{d\theta}{dx} \right)
\]
\[
\Rightarrow \frac{d^2y(x)}{dx^2} = -\frac{\cos(\theta)}{\sin^2(\theta)} \frac{d\Theta}{d\theta} + \frac{1}{\sin^2(\theta)} \frac{d^2\Theta}{d\theta^2}
\] (12.197)
where we used the product rule. Then using the results above the associated Legendre differential equation can be expressed as
\[
\sin^2(\theta) \left[ -\frac{\cos(\theta)}{\sin^3(\theta)} \frac{d\Theta}{d\theta} + \frac{1}{\sin^2(\theta)} \frac{d^2\Theta}{d\theta^2} \right] + 2\frac{\cos(\theta)}{\sin(\theta)} \frac{d\Theta}{d\theta}
\]
\[
+ \left[ l(l+1) - \frac{m^2}{\sin^2(\theta)} \right] \Theta = 0
\] (12.198)
\[
\Rightarrow \frac{d^2\Theta}{d\theta^2} + \frac{\cos(\theta)}{\sin(\theta)} \frac{d\Theta}{d\theta} + \left[ l(l+1) - \frac{m^2}{\sin^2(\theta)} \right] \Theta = 0
\] (12.199)
that can be put in the form
\[
\frac{1}{\sin(\theta)} \frac{d}{d\theta} \left[ \sin(\theta) \frac{d\Theta}{d\theta} \right] + \left[ l(l+1) - \frac{m^2}{\sin^2(\theta)} \right] \Theta = 0.
\] (12.200)

The solution to this form of the associated Legendre differential equation can then be written as
\[
\Theta(\theta) = P_l^m(\cos(\theta)) = \sin^m(\theta) \frac{d^m}{d(\cos(\theta))^m} P_l(\cos(\theta)).
\] (12.201)

The set of functions defined by
\[
Y_{lm}(\theta, \varphi) = C(l, m) P_l^m(\cos(\theta)) e^{im\varphi},
\]
are known as Spherical harmonics, where
\[
C(l, m) = (-1)^m \sqrt{\frac{2l+1 (l-m)!}{4\pi (l+m)!}},
\]
are determined from the normalization condition
\[
\int_0^\pi \int_0^{2\pi} Y_{lm}^*(\theta, \varphi) Y_{lm}(\theta, \varphi) \sin(\theta) \, d\theta \, d\varphi = 1.
\]
Note that
\[ l = 0, 1, 2, 3... \] (12.202)
and
\[ m^2 \leq l^2 \Rightarrow m = -l, -l + 1, -l + 2...0, 1, 3...l. \]
The Spherical harmonics form an orthonormal set of vectors/functions and the orthonormality condition is given by
\[
\int_0^{2\pi} \int_0^\pi Y_{lm}^* (\theta', \varphi') Y_{l'm'} (\theta, \varphi) \sin (\theta) d\theta d\varphi = \delta_{ll'}\delta_{mm'}. \tag{12.203}
\]
Spherical harmonics form a complete set of vectors, which we may express, use Dirac’s notion, as
\[
\sum_{l=0}^{\infty} \sum_{m=-l}^{l} |Y_{lm}\rangle \langle Y_{lm}| = 1
\]
or form a complete set of functions,
\[
\sum_{l=0}^{\infty} \sum_{m=-l}^{l} \int_0^{2\pi} \int_0^\pi \langle \theta', \varphi' | Y_{lm}\rangle \langle Y_{lm}| \theta', \varphi' \rangle \sin (\theta') d\theta' d\varphi' = 1
\]
\[
\Rightarrow \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \int_0^{2\pi} \int_0^\pi Y_{lm}^* (\theta', \varphi') Y_{lm} (\theta', \varphi') \sin (\theta') d\theta' d\varphi' = 1.
\]
Thus any finite function, \( f (\theta, \varphi) \)
\[
f (\theta, \varphi) = \langle \theta, \varphi | f \rangle
\]
that is defined in the domain \( \theta \in [0, \pi] \) and \( \varphi \in [0, 2\pi] \) can be expressed as linear combination of the spherical harmonics
\[
f (\theta, \varphi) = \langle \theta, \varphi | f \rangle = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \langle \theta, \varphi | Y_{lm}\rangle \langle Y_{lm}| f \rangle \tag{12.204}
\]
Using the completeness relation for the continuos angular space
\[
\int_0^{2\pi} \int_0^\pi |\theta', \varphi'\rangle \langle \theta', \varphi' | \sin (\theta') d\theta' d\varphi' = 1, \tag{12.205}
\]
where
\[
\langle \theta | \theta' \rangle = \sigma (\theta - \theta'), \langle \varphi | \varphi' \rangle = \sigma (\varphi - \varphi') \tag{12.206}
\]
we can rewrite
\[
f (\theta, \varphi) = \langle \theta, \varphi | f \rangle = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \langle \theta, \varphi | Y_{lm}\rangle \langle Y_{lm}| f \rangle \tag{12.207}
\]
CHAPTER 12. POWER SERIES SOLUTIONS TO DIFFERENTIAL EQUATIONS

as

\[ f(\theta, \varphi) = \langle \theta, \varphi \mid f \rangle \]

\[ = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \int_{0}^{2\pi} \int_{0}^{\pi} \langle \theta, \varphi \mid Y_{lm} \rangle \langle Y_{lm} \mid \theta', \varphi' \rangle \langle \theta', \varphi' \mid f \rangle \sin(\theta') \, d\theta' \, d\varphi' \]

\[ = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \int_{0}^{2\pi} \int_{0}^{\pi} \langle \theta, \varphi \mid Y_{lm} \rangle \langle Y_{lm} \mid \theta', \varphi' \rangle \langle \theta', \varphi' \mid f \rangle \sin(\theta') \, d\theta' \, d\varphi' \]  \hspace{1cm} (12.208)

Using the Dirac’s notation

\[ Y_{lm}(\theta, \varphi) = \langle \theta, \varphi \mid Y_{lm} \rangle, Y_{lm}^*(\theta', \varphi') = \langle Y_{lm} \mid \theta', \varphi' \rangle, \]

\[ f(\theta', \varphi') = \langle \theta', \varphi' \mid f \rangle, \]  \hspace{1cm} (12.209)

one can write

\[ f(\theta, \varphi) = \langle \theta, \varphi \mid f \rangle \]

\[ = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left[ \int_{0}^{2\pi} \int_{0}^{\pi} Y_{lm}^*(\theta', \varphi') \, f(\theta', \varphi') \sin(\theta') \, d\theta' \, d\varphi' \right] Y_{lm}(\theta, \varphi) \]  \hspace{1cm} (12.210)

or

\[ f(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm} Y_{lm}(\theta, \varphi), \]  \hspace{1cm} (12.211)

where

\[ A_{lm} = \int_{0}^{\pi} \int_{0}^{2\pi} Y_{lm}^*(\theta, \varphi) \, f(\theta, \varphi) \sin(\theta) \, d\theta d\varphi. \]  \hspace{1cm} (12.212)

are the expansion coefficients.

The Inverse-Distance Between Two Points: Consider two points described by the vectors, \( \vec{r} \) and \( \vec{r}' \), as shown in Fig. 12.3. In spherical coordinates these two points are described by \( (r, \theta, \varphi) \) and \( (r', \theta', \varphi') \), respectively. The angle between these two vectors, \( \gamma \), (which is related to the angles describing each of the two vectors) can be used to determine the inverse-distance between these two points. In view of the equation we determined that relates the inverse distance and the Legendre polynomials, one can write

\[ \frac{1}{|\vec{r} - \vec{r}'|} = \frac{1}{\sqrt{r^2 + r'^2 - 2rr' \cos(\gamma)}} = \sum_{l=0}^{\infty} r'^{l+1} P_l(\cos(\gamma)) \]  \hspace{1cm} (12.213)

where

\[ \frac{r'}{r} = \begin{cases} \frac{r'r'}{r^2}, & \text{for } r' < r \\ \frac{r'^2}{r^2}, & \text{for } r < r' \end{cases} \]
Figure 12.3: Two point described by two vectors, \( \mathbf{r} \) and \( \mathbf{r}' \). In spherical coordinates the two points have coordinates \((r, \theta, \varphi)\) and \((r', \theta', \varphi')\), respectively. The angle between these two vectors is \( \gamma \).

The Addition Theorem for Spherical Harmonics: A mathematical result of considerable interest and the use is called the addition theorem for spherical Harmonics. The addition theorem states that

\[
P_l(\cos(\gamma)) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_{lm}(\theta, \varphi) Y_{lm}^*(\theta', \varphi'),
\]

where \( P_l \) are the Legendre polynomials and \( \cos(\gamma) \)

\[
\cos(\gamma) = \cos(\theta) \cos(\theta') + \sin(\theta) \sin(\theta') \cos(\varphi - \varphi').
\]

We can prove this theorem using

\[
\mathbf{r} = r \sin(\theta) \cos(\varphi) \hat{x} + r \sin(\theta) \sin(\varphi) \hat{y} + r \cos(\theta) \hat{z},
\]

and

\[
\mathbf{r}' = r' \sin(\theta') \cos(\varphi') \hat{x} + r' \sin(\theta') \sin(\varphi') \hat{y} + r' \cos(\theta') \hat{z}.
\]

We note that

\[
|\mathbf{r} - \mathbf{r}'| = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}
\]

\[
= \left\{ (r \sin(\theta) \cos(\varphi) - r' \sin(\theta') \cos(\varphi'))^2 \ight. \\
+ \left. (r \sin(\theta) \sin(\varphi) - r' \sin(\theta') \sin(\varphi'))^2 + (r \cos(\theta) - r' \cos(\theta'))^2 \right\}
\]
which can be rewritten as
\[
|\vec{r} - \vec{r}'| = r^2 \sin^2 (\theta) \cos^2 (\varphi) + r'^2 \sin^2 (\theta') \cos^2 (\varphi') \\
- 2rr' \sin (\theta) \cos (\varphi) \cos (\varphi') + r^2 \sin^2 (\theta) \sin (\varphi) + r'^2 \sin^2 (\theta') \sin^2 (\varphi') \\
- 2rr' \sin (\theta) \sin (\theta') \sin (\varphi) \sin (\varphi') + r^2 \cos^2 (\theta) + r'^2 \cos^2 (\theta') \\
- 2rr' \cos (\theta) \cos (\theta')
\] (12.219)

Applying the relations
\[
\cos (\varphi - \varphi') = \cos (\varphi) \cos (\varphi') + \sin (\varphi) \sin (\varphi') \\
\cos^2 (\varphi) + \sin^2 (\varphi) = 1, \sin^2 (\theta) + \cos^2 (\theta) = 1
\] (12.221)

one finds
\[
|\vec{r} - \vec{r}'| = r^2 + r'^2 - 2rr' \left[ \cos (\theta) \cos (\theta') + \sin (\theta) \sin (\theta') \cos (\varphi - \varphi') \right].
\] (12.222)

This can be put in the form
\[
|\vec{r} - \vec{r}'| = r^2 + r'^2 - 2rr' \cos (\gamma),
\] (12.223)

where
\[
\cos (\gamma) = \cos (\theta) \cos (\theta') + \sin (\theta) \sin (\theta') \cos (\varphi - \varphi')
\] (12.224)

and \(\gamma\) is the angle between the two vectors, \(\vec{r}\) and \(\vec{r}'\), shown in Fig. 12.3.

**Example 12.8** A solid sphere of radius \(R\) has a constant volume-charge density \(\rho\). The sphere is centered at the origin of coordinates. The point \(P\) is a distance \(d > R\) from the center of the sphere. Find an expression for the electrostatic potential at the point \(P\), \(V_P\), due to the charged sphere.

**Solution:** The electrostatic potential at point \(P\) is given by the integral expression
\[
V (\vec{r}) = \frac{1}{4 \pi \epsilon_0} \int_{\text{vol}} \frac{dq'_{\text{vol}}}{|\vec{r} - \vec{r}'|}.
\] (12.225)

For a uniform charge density, \(\rho\), we have
\[
\frac{dq'}{dt'} = \rho \Rightarrow dq' = \rho dt' = \rho r'^2 dr' \sin (\theta') d\theta' d\varphi'
\] (12.226)
Figure 12.4: A charged solid sphere with uniform charge density, \( \rho \).

where we used spherical coordinates for the infinitesimal value \( dr' \). Using this relation the potential can then be expressed as

\[
V ( \vec{r} ) = \frac{\rho}{4\pi\varepsilon_0} \int_0^R \int_0^{\pi} \int_0^{2\pi} \frac{1}{|r' - r|^2} r'^2 dr' \sin (\theta') d\theta' d\varphi'.
\] (12.227)

The relation

\[
\frac{1}{|r' - r|^2} = \sum_{l=0}^\infty \sum_{m=-l}^l \frac{r'^l}{r^{l+1}} P_l (\cos \gamma),
\] (12.228)

for a point outside the sphere (\( r = d > r' \)) give

\[
\frac{1}{|r' - r|^2} = \sum_{l=0}^\infty \sum_{m=-l}^l \frac{r'^l}{r^{l+1}} P_l (\cos \gamma).
\] (12.229)

and the potential becomes

\[
V ( \vec{r} ) = \frac{\rho}{4\pi\varepsilon_0} \sum_{l=0}^\infty \frac{r'^l}{d^{l+1}} \int_0^R \int_0^{\pi} \int_0^{2\pi} P_l (\cos \gamma) r'^2 dr' \sin (\theta') d\theta' d\varphi'.
\] (12.230)

Now applying the addition theorem for the spherical harmonics

\[
P_l (\cos \gamma) = \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_{lm} (\theta, \varphi) Y_{lm}^* (\theta', \varphi')
\] (12.231)
we find

\[ V(\vec{r}) = \frac{\rho}{4\pi \epsilon_0} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left( \frac{4\pi}{2l+1} \right) \]

\[ \times \int_0^R \int_\theta^\pi \int_{\varphi=0}^{2\pi} \frac{r^l}{l!+1} Y_{lm}(\theta, \varphi) Y^{*}_{lm}(\theta', \varphi') r'^2 d\theta' \sin(\theta') d\theta' d\varphi'. \]

(12.232)

Since the integration is with respect to the prime variables we can rewrite the above expression in the form

\[ V(\vec{r}) = \frac{\rho}{\epsilon_0} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{r^l}{l!+1} \left( \frac{1}{2l+1} \right) Y_{lm}(\theta, \varphi) \]

\[ \times \int_0^R \frac{r'^{l+2}}{l!+1} dr' \int_\theta^\pi \int_{\varphi=0}^{2\pi} Y^{*}_{lm}(\theta', \varphi') \sin(\theta') d\theta' d\varphi'. \]

(12.233)

We note that for, \( l = m = 0 \), the relations,

\[ Y_{lm}(\theta', \varphi') = (-1)^m \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta') e^{im\varphi'} \]

(12.234)

gives

\[ Y_{00}(\theta, \varphi) = \sqrt{\frac{1}{4\pi}} \Rightarrow \sqrt{4\pi} Y_{00}(\theta, \varphi) = 1. \]

(12.235)

Substituting this into the expression for the potential, one can then

\[ V(\vec{r}) = \frac{\rho}{\epsilon_0} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left( \frac{1}{2l+1} \right) Y_{lm}(\theta, \varphi) \]

\[ \times \int_0^R \frac{r'^{l+2}}{l!+1} dr' \int_\theta^\pi \int_{\varphi=0}^{2\pi} Y^{*}_{lm}(\theta', \varphi') Y_{lm}(\theta', \varphi') \sin(\theta') d\theta' d\varphi'. \]

(12.236)

Now using the orthonormality relation for the spherical harmonics,

\[ \langle Y_{lm} | Y_{l'm'} \rangle = \int_0^\pi \int_{\varphi=0}^{2\pi} Y^*_{lm}(\theta', \varphi') Y_{l'm'}(\theta', \varphi') \sin(\theta') d\theta' d\varphi' \]

\[ = \delta_{ll'} \delta_{mm'} \]

(12.237)

we have

\[ \int_\theta^\pi \int_{\varphi=0}^{2\pi} Y_{00}(\theta', \varphi') Y^{*}_{lm}(\theta', \varphi') \sin(\theta') d\theta' d\varphi' = \delta_{0l} \delta_{0m} \]

(12.238)
so that the potential becomes

\[ V(\vec{r}) = \frac{\rho \sqrt{4\pi}}{4\pi \varepsilon_0} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left( \frac{4\pi}{2l+1} \right) \delta_{0l} \delta_{0m} Y_{lm}(\theta, \varphi) \int_0^R \frac{r^{l+2}}{d^{l+1}} dr'. \]  

(12.239)

All the terms in the summation become zero except when \( l = 0 \) and \( m = 0 \) because of the the Kronecker delta function. Thus we find for the potential

\[ V(\vec{r}) = \frac{\rho \sqrt{4\pi}}{4\pi \varepsilon_0} \frac{Y_{00}(\theta, \varphi)}{d} \int_0^R r^2 dr' = \frac{\rho \sqrt{4\pi}}{4\pi \varepsilon_0} \frac{Y_{00}(\theta, \varphi) R^3}{3d}, \]  

\[ \Rightarrow V(\vec{r}) = \frac{\rho R^3}{3\varepsilon_0 d}. \]  

(12.240)

In terms of the total charge, \( Q \), for a uniform distribution the charge density is given by

\[ \frac{Q}{\frac{4}{3}\pi R^3} = \rho, \]  

(12.241)

and the potential becomes

\[ V(\vec{r}) = \frac{Q}{4\pi \varepsilon_0 d}, \]  

(12.242)

which is the potential of a point charge.

### 12.10 The Method of Frobenius and the Bessel equation

**When the Standard Power Series Solutions fail**: The power series solution

\[ y(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \ldots, \]  

(12.243)

where \( n \geq 0 \). This is applicable when the differential equation involves no singular point for \( \forall x \in \mathbb{R} \). There are differential equations which are not defined for \( \forall x \in \mathbb{R} \). The solutions of such differential equations involves a factor with negative or fractions as exponent for \( x \). In such cases the standard power series solution fails and we need to use a slightly modified method known *The Method of Frobenius* which we will see next. But before we do so we first consider some simple differential equations where the solutions do not satisfy the standard power series solution.

**Example 12.9** Solve the differential equation

\[ y' + \frac{y}{x} = 0 \]  

(12.244)
CHAPTER 12. POWER SERIES SOLUTIONS TO DIFFERENTIAL EQUATIONS

Solution: First we note that this differential equation is not defined at \( x = 0 \).
But it can be solved easily as follows:

\[
y' + \frac{y}{x} = 0 \Rightarrow \frac{dy}{dx} = -\frac{1}{x}y \Rightarrow \frac{dy}{y} = -\frac{dx}{x}
\]

\[
\Rightarrow \int_{y_0}^{y} \frac{dy}{y} = - \int_{x_0}^{x} \frac{dx}{x} \Rightarrow \ln \left( \frac{y}{y_0} \right) = - \ln \left( \frac{x}{x_0} \right) = \ln \left( \frac{x}{x_0} \right)^{-1}
\]

\[
\Rightarrow \ln \left( \frac{y}{y_0} \right) = \ln \left( \frac{x_0}{x} \right) \Rightarrow \frac{y}{y_0} = \frac{x_0}{x}
\]

\[
\Rightarrow y = \left( x_0y_0 \right) x^{-1} \quad (12.245)
\]

Noting that \( y_0x_0 = a_1 \) (constant), the solution can be written as

\[
y = a_1 x^{-1}, \quad (12.246)
\]

which clearly show that \( x \) has a negative exponent and the standard power series expansion does not work!

Example 12.10 Solve the differential equation

\[
y' - \frac{3y}{2x} = 0 \quad (12.247)
\]

Solution: Here also the differential equation is singular at \( x = 0 \). The solution is given by

\[
y' - \frac{3y}{2x} = 0 \Rightarrow \frac{dy}{dx} = \frac{3}{2} \frac{1}{x}y \Rightarrow \frac{dy}{y} = \frac{3}{2} \frac{dx}{x} \Rightarrow \int_{y_0}^{y} \frac{dy}{y} = \frac{3}{2} \int_{x_0}^{x} \frac{dx}{x}
\]

\[
\Rightarrow \ln \left( \frac{y}{y_0} \right) = \frac{3}{2} \ln \left( \frac{x}{x_0} \right) = \ln \left( \frac{x}{x_0} \right)^{3/2} \Rightarrow \frac{y}{y_0} = \left( \frac{x}{x_0} \right)^{3/2}
\]

\[
\Rightarrow y = \left( y_0 x_0^{3/2} \right) x^{3/2} \quad (12.248)
\]

Noting that \( y_0 x_0^{3/2} = a_1 \) (constant), the solution can be written as

\[
y = a_1 x^{3/2} = x^{1/2} \left( a_1 x \right). \quad (12.249)
\]

which also involve \( x \) with a fraction exponent.

The Method of Frobenius: When we face a differential equation which are not defined for all \( x \in \mathbb{R} \) like the examples we saw above, we use the method of Frobenius. We used a generalized power series of the form

\[
y(x) = x^s \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{n+s}, \quad (12.250)
\]

where \( s \) is a number to be determined along with the expansion coefficients, \( a_n \).
It may be positive, negative, fraction, or even complex number although we do not consider the complex number case here.
Example 12.11  The Bessel Differential equation: As an application of the Method of Frobenius we will solve the Bessel differential equation

\[ x^2 y'' + xy' + (x^2 - p^2) y = 0, \quad (12.251) \]

where \( p \) is a constant parameter characterizing the differential equation. The Bessel differential equation, usually, derived from the Laplace equation in cylindrical coordinates which we will see in the next chapter.

Solution: Using a generalized power series

\[ y(x) = x^s \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n x^{n+s} \quad (12.252) \]

we have

\[ (x^2 - p^2) y = \sum_{n=0}^{\infty} a_n x^{n+s+2} - \sum_{n=0}^{\infty} p^2 a_n x^{n+s}, \quad (12.253) \]

\[ y'(x) = \sum_{n=0}^{\infty} (n + s) a_n x^{n+s-1} \Rightarrow xy'(x) = \sum_{n=0}^{\infty} (n + s) a_n x^{n+s}, \quad (12.254) \]

\[ y''(x) = \sum_{n=0}^{\infty} (n + s)(n + s - 1) a_n x^{n+s-2} \]

\[ \Rightarrow x^2 y''(x) = \sum_{n=0}^{\infty} (n + s)(n + s - 1) a_n x^{n+s}, \quad (12.255) \]

so that substituting these expressions into the Bessel differential equation, we find

\[ x^2 y'' + xy' + (x^2 - p^2) y = 0 \]

\[ \Rightarrow \sum_{n=0}^{\infty} (n + s)(n + s - 1) a_n x^{n+s} + \sum_{n=0}^{\infty} (n + s) a_n x^{n+s} + \sum_{n=0}^{\infty} a_n x^{n+s+2} \]

\[ -p^2 \sum_{n=0}^{\infty} a_n x^{n+s} = 0 \Rightarrow \sum_{n=0}^{\infty} \left\{ (n + s)(n + s - 1) + (n + s) - p^2 \right\} a_n x^{n+s} \]

\[ + \sum_{n=0}^{\infty} a_n x^{n+s+2} = 0 \Rightarrow \sum_{n=0}^{\infty} \left[ (n + s)^2 - p^2 \right] a_n x^{n+s} + \sum_{n=0}^{\infty} a_n x^{n+s+2} = 0 \]

(12.256)
Expanding the first series in the above expression

\[
\sum_{n=0}^{\infty} \left[ (n + s)^2 - p^2 \right] a_n x^{n+s} = \left[ (0 + s)^2 - p^2 \right] a_0 x^s + \left[ (1 + s)^2 - p^2 \right] a_1 x^{s+1} + \sum_{n=2}^{\infty} \left[ (n + s)^2 - p^2 \right] a_n x^{s+n}
\]

\[
= [s^2 - p^2] a_0 x^s + [(1 + s)^2 - p^2] a_1 x^{s+1} + \sum_{n=2}^{\infty} [(n + s)^2 - p^2] a_n x^{n+s+1} = 0 \quad \text{(12.257)}
\]

and replacing \( n = m + 2 \) and noting that for \( n = 2 \Rightarrow m = 0 \), we can write

\[
\sum_{n=0}^{\infty} \left[ (n + s)^2 - p^2 \right] a_n x^{n+s} = \left[ s^2 - p^2 \right] a_0 x^s + \left[ (1 + s)^2 - p^2 \right] a_1 x^{s+1} + \sum_{m=0}^{\infty} \left[ (m + 2 + s)^2 - p^2 \right] a_{m+2} x^{m+s+2} = 0 \quad \text{(12.258)}
\]

Since \( m \) is a dummy variable we can replace it by \( n \) and rewrite the above expression as

\[
\sum_{n=0}^{\infty} \left[ (n + s)^2 - p^2 \right] a_n x^{n+s} = \left[ s^2 - p^2 \right] a_0 x^s + \left[ (1 + s)^2 - p^2 \right] a_1 x^{s+1} + \sum_{n=0}^{\infty} \left[ (n + 2 + s)^2 - p^2 \right] a_{n+2} x^{n+s+2} = 0 \quad \text{(12.259)}
\]

Now substituting this expression into

\[
\sum_{n=0}^{\infty} \left[ (n + s)^2 - p^2 \right] a_n x^{n+s} + \sum_{n=0}^{\infty} a_n x^{n+s+2} = 0 \quad \text{(12.260)}
\]

we find

\[
\left[ s^2 - p^2 \right] a_0 x^s + \left[ (1 + s)^2 - p^2 \right] a_1 x^{s+1} + \sum_{n=0}^{\infty} \left[ (n + 2 + s)^2 - p^2 \right] a_{n+2} x^{n+s+2} + \sum_{n=0}^{\infty} a_n x^{n+s+2} = 0 \quad \text{(12.261)}
\]

\[
\left[ s^2 - p^2 \right] a_0 x^s + \left[ (1 + s)^2 - p^2 \right] a_1 x^{s+1} + \sum_{n=0}^{\infty} \left( \left[ (n + 2 + s)^2 - p^2 \right] a_{n+2} + a_n \right) x^{n+s+2} = 0. \quad \text{(12.262)}
\]
There follows that
\[
\begin{align*}
(s^2 - p^2) a_0 x^s &= 0, \\
(1 + s)^2 - p^2) a_1 x^{s+1} &= 0,
\end{align*}
\]
\[\sum_{n=0}^{\infty} \left\{ \left[ (n + 2 + s)^2 - p^2 \right] a_{n+2} + a_n \right\} x^{n+s+2} = 0, \quad (12.263)
\]
which leads to
\[
\begin{align*}
(s^2 - p^2) a_0 &= 0, \\
(1 + s)^2 - p^2) a_1 &= 0, \\
(n + 2 + s)^2 - p^2) a_{n+2} + a_n &= 0
\end{align*}
\]
\[\quad (12.264)
\]
Upon solving the first equation, we find
\[s = \pm p \quad (12.265)\]
and substituting this value into the second equation
\[
\begin{align*}
(1 + s)^2 - p^2) a_1 &= 0
\end{align*}
\]
\[\quad (12.266)\]
we find
\[
\begin{align*}
(1 \pm p)^2 - p^2) a_1 &= 0 \Rightarrow [1 \pm 2p] a_1 = 0.
\end{align*}
\]
\[\quad (12.267)\]
For this to be true independent of \(p\), we must have
\[a_1 = 0. \quad (12.268)\]
The third equation
\[
\begin{align*}
(n + 2 + s)^2 - p^2) a_{n+2} + a_n &= 0
\end{align*}
\]
leads to the recursion relation
\[a_{n+2} = -\frac{1}{(n + 2 + s)^2 - p^2} a_n. \quad (12.270)\]
This recursion relation gives two different functions for the two different values of \(s (= \pm p)\).

First solution \((s = p)\): For this case the recursion formula becomes
\[
a_{n+2} = -\frac{a_n}{(n + 2 + p)^2 - p^2} = -\frac{a_n}{(n + 2)^2 + 2(n + 2)p} = -\frac{a_n}{(n + 2)(n + 2 + 2p)} \quad (12.271)
\]
Since we found \(a_1 = 0\) all the odd terms vanish and the recursion relation can be expressed, for the even terms using \(n + 2 = 2m \Rightarrow n = 2m - 2\), as
\[a_{2m} = -\frac{a_{2m-2}}{(2m)(2m + 2p)} = -\frac{a_{2m-2}}{2^2m(m + p)} \quad (12.272)\]
where \( m = 1, 2, 3 \). Next we shall consider the first few terms:

\[
\begin{align*}
  m &= 1 \implies a_2 = -\frac{a_0}{2^2! (1 + p)} \\
  m &= 2 \implies a_4 = -\frac{a_0}{2^4! (2 + p)} = -\frac{a_0}{2^4! (1 + p) (2 + p)} \\
  m &= 3 \implies a_6 = -\frac{a_0}{2^6! (3 + p)} = -\frac{a_0}{2^6! (1 + p) (2 + p) (3 + p)} \quad (12.273)
\end{align*}
\]

Recalling that the Gamma function

\[
\Gamma(p + 1) = p! \Gamma(p), \quad \Gamma(p + 2) = (p + 1) \Gamma(p + 1) = p (p + 1) \Gamma(p), \quad (12.274)
\]

\[
\Gamma(p + 3) = (p + 2) \Gamma(p + 2) = p (p + 1) (p + 2) \Gamma(p), \quad (12.275)
\]

we can rewrite

\[
\begin{align*}
  m &= 1 \implies a_2 = -\frac{a_0}{2^2! (1 + p)} = -\frac{a_0 p \Gamma(p)}{2^2! p (1 + p) \Gamma(p)} = -\frac{a_0 \Gamma(p + 1)}{2^2! \Gamma(p + 2)} \quad (12.276) \\
  m &= 2 \implies a_4 = -\frac{a_0}{2^4! (1 + p) (2 + p)} = -\frac{a_0 p \Gamma(p)}{2^4! p (1 + p) (2 + p) \Gamma(p)} \implies a_4 = -\frac{a_0 \Gamma(p + 1)}{2^4! \Gamma(p + 3)} \quad (12.277) \\
  m &= 3 \implies a_6 = -\frac{a_0}{2^6! (1 + p) (2 + p) (3 + p)} = -\frac{a_0 p \Gamma(p)}{2^6! p (1 + p) (2 + p) (3 + p) \Gamma(p)} \implies a_6 = -\frac{a_0 \Gamma(p + 1)}{2^6! \Gamma(p + 4)} \quad (12.278)
\end{align*}
\]

Therefore, the solution becomes

\[
y(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{m=0}^{\infty} a_{2m} x^{2m} \\
y(x) = x^p \left\{ a_0 - \frac{a_0 \Gamma(p + 1)}{2^2! \Gamma(p + 2)} x^2 + \frac{a_0 \Gamma(p + 1)}{2^4! \Gamma(p + 3)} x^4 - \frac{a_0 \Gamma(p + 1)}{2^6! \Gamma(p + 4)} x^4 + \ldots \right\} \quad (12.279)
\]

which can be simplified into

\[
y(x) = a_0 x^p \Gamma(p + 1) \left\{ \frac{1}{\Gamma(p + 1)} - \frac{1}{1! \Gamma(p + 2)} \left( \frac{x}{2} \right)^2 + \frac{1}{2! \Gamma(p + 3)} \left( \frac{x}{2} \right)^4 \right. \\
\left. - \frac{1}{3! \Gamma(p + 4)} \left( \frac{x}{2} \right)^6 + \ldots \right\} \quad (12.280)
\]

Noting that

\[
\Gamma(1) = \Gamma(2) = 1, \quad n! = \Gamma(n + 1) \quad (12.281)
\]
the above expression can be rewritten as
\[ y(x) = a_0 2^p \left( \frac{x}{2} \right)^p \Gamma(p + 1) \left\{ \frac{1}{\Gamma(1) \Gamma(p + 1)} - \frac{1}{\Gamma(2) \Gamma(p + 2)} \left( \frac{x}{2} \right)^2 + \frac{1}{\Gamma(3) \Gamma(p + 3)} \left( \frac{x}{2} \right)^4 - \frac{1}{\Gamma(4) \Gamma(p + 4)} \left( \frac{x}{2} \right)^6 + \ldots \right\}. \]

Introducing the function, \( J_p(x) \), defined by
\[ J_p(x) = \frac{y(x)}{2^p a_0 \Gamma(p + 1)}, \] (12.282)
we find
\[ J_p(x) = \left( \frac{x}{2} \right)^p \left\{ \frac{1}{\Gamma(1) \Gamma(p + 1)} - \frac{1}{\Gamma(2) \Gamma(p + 2)} \left( \frac{x}{2} \right)^2 + \frac{1}{\Gamma(3) \Gamma(p + 3)} \left( \frac{x}{2} \right)^4 - \frac{1}{\Gamma(4) \Gamma(p + 4)} \left( \frac{x}{2} \right)^6 + \ldots \right\}. \] (12.283)
that can be put in the form
\[ J_p(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n + 1) \Gamma(n + 1 + p)} \left( \frac{x}{2} \right)^{2n+p}. \] (12.284)

This function \( J_p(x) \) is called the Bessel function of the first kind of order \( p \).

**Second Solution \( s = -p \):** The second solution, when \( s = -p \) can easily be obtained from the first solution by replacing \( p \) with \(-p\).
\[ J_{-p}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n + 1) \Gamma(n + 1 - p)} \left( \frac{x}{2} \right)^{2n-p}. \] (12.285)

If \( p \) is not an integer, \( J_p(x) \) is a series starting with \( x^p \) and \( J_{-p}(x) \) is a series starting with \( x^{-p} \). Then \( J_p(x) \) and \( J_{-p}(x) \) are two independent solutions and the linear combination of these two functions is a general solution. The combination is is called either the Neumann or Weber function denoted by either \( N_p(x) \) or \( Y_p(x) \) and is given by
\[ N_p(x) = Y_p(x) = \frac{\cos(\pi p) J_p(x) - J_{-p}(x)}{\sin(\pi p)}. \] (12.286)
However, if \( p \) is an integer, then the first few terms in \( J_{-p}(x) \) are zero because \( \Gamma(n + 1 - p) \) is the gamma function of a negative integer, which is infinite.
\[ J_{-p}(x) = (-1)^p J_p(x) \] (12.287)
and \( J_p(x) \) and \( J_{-p}(x) \) are not independent solutions.

**The zeroes of the Bessel function:** The zeroes of the bessel function \( \{ \alpha_1, \alpha_2, \alpha_3, \ldots \} \) are the values of \( x \) at which
\[ J_p(x) = 0. \] (12.288)
in the figure below the points on the \( x \) axis where the \( J_p(x) \) intersects (i.e. \( J_p(x) = 0 \)) are the zeroes of the Bessel function.
12.11 The orthogonality of the Bessel Functions

The Orthogonality and Normalization: In the previous section we did solve the Bessel Differential Equation

\[ x^2 y'' + xy' + (x^2 - p^2) y = 0 \]  \hspace{1cm} (12.289)

and have found the solution is the Bessel function \( J_p(x) \). We now determine the orthonormality of these functions. To this end, noting that

\[ x^2 y'' + xy' = x (xy'' + y') = x \frac{d}{dx} \left( x \frac{dy}{dx} \right) \]

we can rewrite the Bessel Equation as

\[ x \frac{d}{dx} \left( x \frac{dy}{dx} \right) + (x^2 - p^2) y = 0 \]  \hspace{1cm} (12.290)

the solution of which is given by the bessel functions, \( y(x) = J_p(x) \). Thus for a the Bessel function, \( J_p(u) \), one can write

\[ u \frac{d}{du} \left( u \frac{dJ_p(u)}{du} \right) + (u^2 - p^2) J_p(u) = 0. \]  \hspace{1cm} (12.292)

We introduce the transformation of variable defined by

\[ u = \alpha x \Rightarrow du = \alpha dx, \]  \hspace{1cm} (12.293)
12.11. THE ORTHOGONALITY OF THE BESSEL FUNCTIONS

where

\[ \alpha_i = \{\alpha_1, \alpha_2, \alpha_3, \ldots\} \quad (12.294) \]

are the zeros of the Bessel function. Then one can write for \( \alpha = \alpha_i \)

\[ x \frac{d}{dx} \left( x \frac{dJ_p(\alpha_i x)}{dx} \right) + \left( \alpha_i^2 x^2 - p^2 \right) J_p(\alpha_i x) = 0, \quad (12.295) \]

and \( \alpha = \alpha_j \)

\[ x \frac{d}{dx} \left( x \frac{dJ_p(\alpha_j x)}{dx} \right) + \left( \alpha_j^2 x^2 - p^2 \right) J_p(\alpha_j x) = 0 \quad (12.296) \]

Multiplying Eq. (12.295) by \( J_p(\alpha_j x) \) and Eq. (12.296) by \(-J_p(\alpha_i x)\), we have

\[ xJ_p(\alpha_j x) \frac{d}{dx} \left( x \frac{dJ_p(\alpha_i x)}{dx} \right) + \left( \alpha_i^2 x^2 - p^2 \right) J_p(\alpha_j x) J_p(\alpha_i x) = 0, \quad (12.297) \]

and

\[ -xJ_p(\alpha_i x) \frac{d}{dx} \left( x \frac{dJ_p(\alpha_j x)}{dx} \right) - \left( \alpha_j^2 x^2 - p^2 \right) J_p(\alpha_i x) J_p(\alpha_j x) = 0 \quad (12.298) \]

so that adding these two equations leads to

\[ J_p(\alpha_j x) \frac{d}{dx} \left( x \frac{dJ_p(\alpha_i x)}{dx} \right) - J_p(\alpha_i x) \frac{d}{dx} \left( x \frac{dJ_p(\alpha_j x)}{dx} \right) + \left( \alpha_i^2 - \alpha_j^2 \right) J_p(\alpha_i x) J_p(\alpha_j x) x = 0. \quad (12.299) \]
Integrating this equation with respect to $x$ from 0 to 1

$$
\int_0^1 J_p(\alpha_j x) \frac{d}{dx} \left( x \frac{d}{dx} J_p(\alpha_i x) \right) dx - \int_0^1 J_p(\alpha_i x) \frac{d}{dx} \left( x \frac{d}{dx} J_p(\alpha_j x) \right) dx
+ (\alpha_i^2 - \alpha_j^2) \int_0^1 J_p(\alpha_i x) J_p(\alpha_j x) x dx = 0.
$$

(12.300)

Applying integration by parts,

$$
\int_0^1 u \frac{dv}{dx} dx = uv\big|_0^1 - \int_0^1 v \frac{du}{dx} dx,
$$

(12.301)

one can write

$$
\int_0^1 J_p(\alpha_j x) \frac{d}{dx} \left( x \frac{d}{dx} J_p(\alpha_i x) \right) dx = J_p(\alpha_j x) x \frac{d}{dx} (J_p(\alpha_i x))\big|_0^1
- \int_0^1 x \frac{d}{dx} (J_p(\alpha_i x)) \frac{d}{dx} (J_p(\alpha_j x)) dx
$$

and

$$
\int_0^1 J_p(\alpha_i x) \frac{d}{dx} \left( x \frac{d}{dx} J_p(\alpha_j x) \right) = J_p(\alpha_i x) x \frac{d}{dx} (J_p(\alpha_j x))\big|_0^1
- \int_0^1 x \frac{d}{dx} (J_p(\alpha_j x)) \frac{d}{dx} (J_p(\alpha_i x)) dx.
$$

(12.302)

(12.303)

Noting that

$$
x = 0 \Rightarrow x J_p(\alpha_i x) = 0,
$$

(12.304)

and $\alpha_i$ & $\alpha_j$ are the zeros of the Bessel function

$$
x = 1 \Rightarrow J_p(\alpha_i x) x = J_p(\alpha_i) = 0,
$$

(12.305)

one finds

$$
\int_0^1 J_p(\alpha_j x) \frac{d}{dx} \left( x \frac{d}{dx} J_p(\alpha_i x) \right) dx = - \int_0^1 x \frac{d}{dx} (J_p(\alpha_i x)) \frac{d}{dx} (J_p(\alpha_j x)) dx
$$

(12.306)

and

$$
\int_0^1 J_p(\alpha_i x) \frac{d}{dx} \left( x \frac{d}{dx} J_p(\alpha_j x) \right) = - \int_0^1 x \frac{d}{dx} (J_p(\alpha_j x)) \frac{d}{dx} (J_p(\alpha_i x)) dx.
$$

(12.307)

Now substituting Eq. (12.306) and Eq. (12.307) into Eq. (12.300), one finds

$$
(\alpha_i^2 - \alpha_j^2) \int_0^1 J_p(\alpha_i x) J_p(\alpha_j x) x dx = 0.
$$

(12.308)
There follows that, clearly, for \( i \neq j \), since
\[
\alpha_i^2 - \alpha_j^2 \neq 0,
\]
we must have
\[
\int_0^1 J_p(\alpha_i x) J_p(\alpha_j x) \, dx = 0. \tag{12.309}
\]
But for \( i = j \), since
\[
\alpha_i^2 - \alpha_j^2 = \alpha_i^2 - \alpha_i^2 = 0, \tag{12.310}
\]
we should generally expect that
\[
\int_0^1 J_p(\alpha_i x) J_p(\alpha_j x) \, dx = \int_0^1 J_p^2(\alpha_i x) \, dx = \text{Constant} \tag{12.311}
\]
Therefore the orthogonality relation for the Bessel functions can be written as
\[
\int_0^1 J_p(\alpha_i x) J_p(\alpha_j x) \, dx = \left\{ \begin{array}{ll}
0, & i \neq j \\
C, & i = j = \delta_{ij}C.
\end{array} \right. \tag{12.312}
\]
To find the constant \( C \), we now consider \( \alpha_i \) (which is not exactly the zeroes of the Bessel function) in Eq. (12.300) is related to \( \alpha_j \) (which is exactly the zeroes of the Bessel function) by
\[
\alpha_i = \alpha_j + \epsilon \Rightarrow \alpha_i^2 - \alpha_j^2 = (\alpha_j + \epsilon)^2 - \alpha_j^2 = 2\alpha_j \epsilon + \epsilon^2 \approx 2\alpha_j \epsilon \tag{12.313}
\]
where, \( \epsilon \ll 1 \), is a small constant which we set to zero to make \( \alpha_i = \alpha_j \). It is important to note that the zeros of the Bessel function
\[
J_p(\alpha_j) = 0, J_p(\alpha_i) = J_p(\alpha_j + \epsilon) \neq 0
\]
but when \( \epsilon \to 0 \),
\[
\lim_{\epsilon \to 0} J_p(\alpha_i) = \lim_{\epsilon \to 0} J_p(\alpha_j + \epsilon) = J_p(\alpha_j) = 0.
\]
Then for \( \alpha_i = \alpha_j + \epsilon \), Eqs. (12.302) \& (12.303) become
\[
\int_0^1 J_p(\alpha_j x) \frac{d}{dx} \left( x \frac{d}{dx} J_p((\alpha_j + \epsilon) x) \right) \, dx = - \int_0^1 x \frac{d}{dx} (J_p((\alpha_j + \epsilon) x)) \frac{d}{dx} (J_p(\alpha_j x)) \, dx \tag{12.314}
\]
and
\[
\int_0^1 J_p((\alpha_j + \epsilon) x) \frac{d}{dx} \left( x \frac{d}{dx} J_p(\alpha_j x) \right) = J_p((\alpha_j + \epsilon) x) \frac{d}{dx} (J_p(\alpha_j x)) \bigg|_0^1
\]
\[
- \int_0^1 x \frac{d}{dx} (J_p(\alpha_j x)) \frac{d}{dx} (J_p((\alpha_j + \epsilon) x)) \, dx
\]
\[
= J_p((\alpha_j + \epsilon)) \left[ \frac{d}{dx} (J_p(\alpha_j x)) \right]_{x=1} - \int_0^1 x \frac{d}{dx} (J_p(\alpha_j x)) \frac{d}{dx} (J_p((\alpha_j + \epsilon) x)) \, dx. \tag{12.315}
\]
Now substituting Eqs. (12.313), (12.314), and (12.315) into Eq. (12.300), we find

\[-J_p((\alpha_j + \epsilon)) \left[ \frac{d}{dx} (J_p(\alpha_j x)) \right]_{x=1} + 2\alpha_j \epsilon \int_0^1 J_p((\alpha_j + \epsilon) x) J_p(\alpha_j x) x \, dx = 0.\]

Noting that

\[\frac{d}{dx} (J_p(\alpha_j x)) \bigg|_{x=1} = \alpha_j \frac{dJ_p(x)}{dx} \bigg|_{x=\alpha_j} = \alpha_j J'_p(\alpha_j) \quad (12.316)\]

we can write

\[\int_0^1 J_p((\alpha_j + \epsilon) x) J_p(\alpha_j x) x \, dx = \frac{\alpha_j J_p((\alpha_j + \epsilon)) J'_p(\alpha_j)}{2\alpha_j}. \quad (12.317)\]

The Taylor series expansion for \( f(\epsilon) \) about \( \epsilon = 0 \), we have

\[f(\epsilon) = f(\epsilon)|_{\epsilon=0} + \epsilon \frac{df(\epsilon)}{d\epsilon} \bigg|_{\epsilon=0} + \frac{\epsilon^2}{2!} \frac{d^2 f(\epsilon)}{d\epsilon^2} \bigg|_{\epsilon=0} + \frac{\epsilon^3}{3!} \frac{d^3 f(\epsilon)}{d\epsilon^3} \bigg|_{\epsilon=0} + \ldots \quad (12.318)\]

then for \( f(\epsilon) = J_p(\alpha_j + \epsilon) \), we have

\[f(\epsilon)|_{\epsilon=0} = J_p(\alpha_j + \epsilon)|_{\epsilon=0} = J_p(\alpha), \]
\[\frac{df(\epsilon)}{d\epsilon} \bigg|_{\epsilon=0} = \frac{dJ_p(\alpha_j + \epsilon)}{d\epsilon} \bigg|_{\epsilon=0} = J'_p(\alpha_j), \quad (12.319)\]

so that

\[J_p(\alpha_j + \epsilon) = J_p(\alpha_j) + \epsilon J'_p(\alpha_j) + \frac{\epsilon^2}{2!} J''_p(\alpha_j) x + \ldots \quad (12.320)\]

Since \( \alpha_j \) is the zero’s of the Bessel function we know \( J_p(\alpha_j) = 0 \), one can then write

\[J_p(\alpha_j + \epsilon) = \epsilon \left[ J'_p(\alpha_j) + \frac{\epsilon}{2!} J''_p(\alpha_j) x + \ldots \right] \quad (12.321)\]

Substituting this into Eq. (12.317), we find

\[\int_0^1 J_p((\alpha_j + \epsilon) x) J_p(\alpha_j x) x \, dx = \frac{\alpha_j J'_p(\alpha_j) \left[ J'_p(\alpha_j) + \frac{\epsilon}{2!} J''_p(\alpha_j) x + \ldots \right]}{2\alpha_j}. \quad (12.322)\]

Then in the limit as \( \epsilon \to 0 \), one finds

\[\int_0^1 J_p(\alpha_j x) J_p(\alpha_j x) x \, dx = \frac{J'_p(\alpha_j)}{2} \frac{J'_p(\alpha_j)}{2}. \quad (12.323)\]

Therefore, the orthogonality condition for the Bessel function can be written as

\[\int_0^1 J_p(\alpha_i x) J_p(\alpha_j x) x \, dx = \frac{J'_p(\alpha_i)}{2} \frac{J'_p(\alpha_j)}{2} \delta_{ij}. \quad (12.324)\]
Limiting (Asymptotic) Forms for the Bessel Functions: The Bessel functions have different simpler approximate expressions for different limiting cases. We recall the Bessel Function

\[ J_p(x) = \left( \frac{x}{2} \right)^p \left\{ \frac{1}{\Gamma(1) \Gamma(p + 1)} - \frac{1}{\Gamma(2) \Gamma(p + 2)} \left( \frac{x}{2} \right)^2 \right. \]

\[ + \frac{1}{\Gamma(3) \Gamma(p + 3)} \left( \frac{x}{2} \right)^4 - \frac{1}{\Gamma(4) \Gamma(p + 4)} \left( \frac{x}{2} \right)^6 + \ldots \} \]  

(12.325)

so that for \( x \ll 1 \), dropping all higher order terms beginning from \( \left( \frac{x}{2} \right)^2 \), we have

\[ J_p(x) \simeq \left( \frac{x}{2} \right)^p \frac{1}{\Gamma(1) \Gamma(p + 1)}. \]  

(12.326)

so that using the relations

\[ \Gamma(1) = 1, \Gamma(p + 1) = p! \]  

(12.327)

one finds

\[ J_p(x) \simeq \frac{1}{p!} \left( \frac{x}{2} \right)^p. \]  

(12.328)

Recalling that

\[ J_{-p}(x) = \left( \frac{x}{2} \right)^{-p} \left\{ \frac{1}{\Gamma(1) \Gamma(1-p)} - \frac{1}{\Gamma(2) \Gamma(2-p)} \left( \frac{x}{2} \right)^2 \right. \]

\[ + \frac{1}{\Gamma(3) \Gamma(3-p)} \left( \frac{x}{2} \right)^4 - \frac{1}{\Gamma(4) \Gamma(4-p)} \left( \frac{x}{2} \right)^6 + \ldots \} \]  

(12.329)

which can be put in the form

\[ J_{-p}(x) = \left( \frac{2}{x} \right)^p \left\{ \frac{1}{\Gamma(1) \Gamma(1-p)} - \frac{1}{\Gamma(2) \Gamma(2-p)} \left( \frac{x}{2} \right)^2 \right. \]

\[ + \frac{1}{\Gamma(3) \Gamma(3-p)} \left( \frac{x}{2} \right)^4 - \frac{1}{\Gamma(4) \Gamma(4-p)} \left( \frac{x}{2} \right)^6 + \ldots \]

\[ + \frac{(-1)^n}{\Gamma(n + 1) \Gamma(n + 1 - p)} \left( \frac{x}{2} \right)^{2n} \} \]  

(12.330)

similarly for \( x \ll 1 \), so that recalling that the Gamma function of a negative number is infinity, the none zero terms are

\[ n + 1 - p = 0 \Rightarrow n = p - 1 \]

which leads to

\[ J_{-p}(x) = \left( \frac{2}{x} \right)^p \left\{ \frac{1}{\Gamma(1) \Gamma(1-p)} - \frac{1}{\Gamma(2) \Gamma(2-p)} \left( \frac{x}{2} \right)^2 \right. \]

\[ + \frac{1}{\Gamma(3) \Gamma(3-p)} \left( \frac{x}{2} \right)^4 - \frac{1}{\Gamma(4) \Gamma(4-p)} \left( \frac{x}{2} \right)^6 + \ldots \]

\[ \ldots + \frac{(-1)^n}{\Gamma(p)} \left( \frac{x}{2} \right)^{2p-1} \} \]  

(12.331)
For $p = 0$, we find
\[ J_{-p}(x) \simeq 1 \]
and for $p > 0$
\[ J_{-p}(x) \simeq \left( \frac{2}{x} \right)^p \frac{1}{\Gamma(p)}, \quad (12.332) \]
Then the Neumann function
\[ N_p(x) = \frac{\cos(\pi p) J_p(x) - J_{-p}(x)}{\sin(\pi p)}. \quad (12.333) \]
becomes for $p = 0$
\[ N_p(x) \simeq \frac{2}{\pi} \ln x. \quad (12.334) \]
and for $p > 0$
\[ N_p(x) \simeq -\frac{(p - 1)!}{\pi} \left( \frac{2}{x} \right)^p \quad (12.335) \]
For $x >> 1$
\[ J_p(x) \simeq \sqrt{\frac{2}{\pi x}} \cos \left( x - \frac{2p + 1}{4} \pi \right), \quad N_p(x) \simeq \sqrt{\frac{2}{\pi x}} \sin \left( x - \frac{2p + 1}{4} \pi \right) \quad (12.336) \]
Note: Many differential equations occur in practice that are not of the standard form of the Bessel differential equation that we saw here but whose solution can be written in terms of the Bessel functions. For example the differential equation
\[ y'' + \frac{1 - 2a}{x} y' + \left[ (bcx^{c-1})^2 + \frac{a^2 - p^2 c^2}{x^2} \right] y = 0 \quad (12.337) \]
has the solution
\[ y = x^a Z_p(bcx), \quad (12.338) \]
where $Z_p$ stands for $J_p$ or $N_p$ or any linear combination, and $a, b, c, p$ are constants.

**Example 12.12 Bessel Functions: An Application (The lengthening Pendulum):** Consider a simple pendulum that consists of a mass $m$ attached to a string of length $l_0$ at the initial time $t = 0$. The length of the pendulum is increasing at a steady rate,
\[ \dot{l} = \frac{dl}{dt} = v. \]
Find the equation of motion and determine the solution for small oscillation.
12.11. THE ORTHOGONALITY OF THE BESSEL FUNCTIONS

Solution: At a given time, $t$, let the length of the pendulum be,

$$l(t) = \sqrt{x(t)^2 + y(t)^2}.$$  

(12.339)

Then the gravitational potential energy can be expressed as

$$U = -mg \left[ l \cos(\theta(t)) - l_0 \right]$$  

(12.340)

and the kinetic energy

$$T = \frac{1}{2} m \left( l^2 \dot{\theta}^2 + \dot{l}^2 \theta^2 \right).$$  

(12.341)

Then the Lagrangian becomes

$$L = T - U = \frac{1}{2} m \left( l^2 \dot{\theta}^2 + \dot{l}^2 \right) + mg \left[ l \cos(\theta) - l_0 \right].$$  

(12.342)

Using Euler-Lagrange equations

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = 0,$$  

(12.343)

we find

$$\frac{d}{dt} \left( ml^2 \dot{\theta} \right) + mgl \sin(\theta) = 0 \Rightarrow \frac{d}{dt} \left( l^2 \dot{\theta} \right) + gl \sin(\theta) = 0.$$  

(12.344)

The length is increasing at a constant rate ($v$), which means

$$\dot{l} = v \Rightarrow dt = \frac{dl}{v}$$  

(12.345)
then transforming the DE from \( t \) to \( l \), we find
\[
v \frac{d}{dl} \left( l^2 v \frac{d\theta}{dl} \right) + gl \sin (\theta) = 0 \Rightarrow l^2 v^2 \frac{d^2 \theta}{dl^2} + 2lv^2 \frac{d\theta}{dl} + gl \sin (\theta) = 0
\]
\[
\Rightarrow \frac{d^2 \theta}{dl^2} + \frac{2}{l} \frac{d\theta}{dl} + \frac{g}{lv^2} \sin (\theta) = 0. \tag{12.346}
\]

For small angle \( \theta \), we have \( \sin (\theta) \approx \theta \) and the DE becomes
\[
\frac{d^2 \theta}{dl^2} + \frac{2}{l} \frac{d\theta}{dl} + \frac{g}{lv^2} \theta = 0. \tag{12.347}
\]

We now consider the DE
\[
\frac{d^2 y}{dx^2} + \frac{1 - 2a}{x} \frac{dy}{dx} + \left[ (bcx^{c-1})^2 + \frac{a^2 - p^2 c^2}{x^2} \right] y = 0 \tag{12.348}
\]
the solution of which is the Bessel function given by
\[
y (x) = x^a Z_p \left( bx^c \right). \tag{12.349}
\]
replacing \( x \) by \( l \) and \( y \) by \( \theta \), we have
\[
\frac{d^2 \theta}{dl^2} + \frac{1 - 2a}{l} \frac{d\theta}{dl} + \left[ (bcl^{c-1})^2 + \frac{a^2 - p^2 c^2}{l^2} \right] \theta = 0 \tag{12.350}
\]
so that
\[
\theta (l) = l^a Z_p \left( bl^c \right). \tag{12.351}
\]
Comparing this with the DE we derived for the lengthening pendulum, we note that
\[
\frac{1 - 2a}{l} = \frac{2}{l}, \tag{12.352}
\]
\[
(bcl^{c-1})^2 + \frac{a^2 - p^2 c^2}{l^2} = \frac{g}{lv^2} \tag{12.353}
\]
There follows that
\[
an = \frac{1}{2}, c = \frac{1}{2}, b = 2 \sqrt{\frac{g}{v^2}}, p^2 = \frac{a^2}{c^2} \Rightarrow p = \pm 1 \tag{12.354}
\]
and one can write
\[
\theta (l) = \frac{1}{\sqrt{l}} Z_1 \left( 2 \sqrt{\frac{gl}{v}} \right), \tag{12.355}
\]
which is the solution to the DE for the lengthening pendulum.
12.12 Fuch’s theorem for 2-nd DE

So far we have seen a second order differential equation of the form

\[ y'' + f(x) y' + g(x) y = 0 \]  \hspace{1cm} (12.356)

can be solved using the standard power series expansion method (e.g. the Legendre equation) or the more general Frobenius method (e.g. the Bessel equation). If you think that the Frobenius method can be used to solve any second order differential equation, then you are wrong. There is a necessary condition that must be satisfied if the differential equation is solvable using Frobenius method. Fuch’s Theorem determines the necessary and sufficient condition to apply the Frobenius Method. It states that for the differential equation

\[ y'' + f(x) y' + g(x) y = 0 \]  \hspace{1cm} (12.357)

the solution can be obtained using the Frobenius method if and only if the functions defined by

\[ F(x) = f(x) y' \]  \hspace{1cm} \[ G(x) = g(x) y \]

can be expressed as power series (i.e. a convergent power series)

\[ F(x) = \sum_{n=0}^{\infty} a_n x^n, \quad G(x) = \sum_{n=0}^{\infty} b_n x^n. \]  \hspace{1cm} (12.358)

When this is the case the solutions to the differential equation consist of either two independent Frobenius series

\[ y_1(x) = \sum_{n=0}^{\infty} A_n x^{n+s}, \quad y_2(x) = \sum_{n=0}^{\infty} B_n x^{n+s} \]  \hspace{1cm} (12.359)

or one solution, \( y_1(x) \), which is a Frobenius series, and the second solution which is given by

\[ y_2(x) = y_1(x) \ln x + y(x), \]  \hspace{1cm} (12.360)

where \( y(x) \) is a Frobenius series. That means

\[ y_1(x) = \sum_{n=0}^{\infty} A_n x^{n+s}, \quad y_2(x) = y_1(x) \ln x + \sum_{n=0}^{\infty} B_n x^{n+s} \]  \hspace{1cm} (12.361)

The second case occurs only when the roots of the indicial equation are equal or differ by an integer.

**Example 12.13** Consider the differential equation

\[ y'' + \frac{1}{x^2} y' - \frac{2}{x^3} y = 0 \]  \hspace{1cm} (12.362)

which has the solution

\[ y(x) = e^{1/x}. \]  \hspace{1cm} (12.363)

Discuss the applicability of a Frobenius-type generalized power-series solution to this differential equation.
Solution: We note that the differential equation has the form
\[ y'' + f(x)y' + g(x)y = 0, \quad (12.364) \]
where
\[ f(x) = \frac{1}{x^2}, g(x) = -\frac{2}{x^3} \quad (12.365) \]
Using the given solution
\[ y(x) = e^{1/x} \quad (12.366) \]
we see that
\[ F(x) = f(x)y' = \frac{1}{x^2} \frac{d}{dx} \left( e^{1/x} \right) = -\frac{1}{x^4} e^{1/x} \quad (12.367) \]
and
\[ G(x) = g(x)y = -\frac{2}{x^3} e^{1/x}. \quad (12.368) \]
These two functions, \( F(x) \) and \( G(x) \), have a singular point at \( x = 0 \) and cannot be expressed as Frobenius series and therefore the Frobenius Method cannot be used to find the solution of the differential equation.
Chapter 13

A partial differential equation

A partial differential equation (PDE) is an equation in some unknown function of more than one variable, say \( f = V(x, y, z, t) \), involving possibly different orders of partial derivatives of that function. The following are some examples of PDEs that would find in different branches of physics.

13.1 PDE in physics

**Gauss’ Law for the Electric Field**: The differential form of Gauss’ law states that the electric field, \( \vec{E}(x, y, z) \) of a volume charge distribution \( \rho(x, y, z) \) satisfies the PDE

\[
\nabla \cdot \vec{E}(x, y, z) = \frac{\rho(x, y, z)}{\epsilon_0}.
\]

(13.1)

If we use

\[
\vec{E}(x, y, z) = E_x(x, y, z) \hat{x} + E_y(x, y, z) \hat{y} + E_z(x, y, z) \hat{z},
\]

\[
\nabla = \frac{\partial}{\partial x} \hat{x} + \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z}
\]

(13.2)

we may write Gauss’ Law as

\[
\frac{\partial E_x(x, y, z)}{\partial x} + \frac{\partial E_y(x, y, z)}{\partial y} + \frac{\partial E_z(x, y, z)}{\partial z} = \frac{\rho(x, y, z)}{\epsilon_0}.
\]

(13.3)

**Poisson’s Equation**: If we express the electric field, \( \vec{E}(x, y, z) \) in terms of the electric potential \( V(x, y, z) \) which is given by

\[
\vec{E}(x, y, z) = -\nabla V(x, y, z)
\]

(13.4)
Gauss’s Law can be expressed as

$$\nabla \cdot [-\nabla V(x, y, z)] = \frac{\rho(x, y, z)}{\varepsilon_0} \quad \Rightarrow \nabla^2 V(x, y, z) = -\frac{\rho(x, y, z)}{\varepsilon_0} \quad (13.5)$$

or

$$\frac{\partial^2 V(x, y, z)}{\partial x^2} + \frac{\partial^2 V(x, y, z)}{\partial y^2} + \frac{\partial^2 V(x, y, z)}{\partial z^2} = -\frac{\rho(x, y, z)}{\varepsilon_0}. \quad (13.6)$$

This PDE is known as Poisson’s equation.

Laplace’s Equation: If the charge density $\rho(x, y, z) = 0$, the Poisson’s equation becomes Laplace equation. It is given by

$$\frac{\partial^2 V(x, y, z)}{\partial x^2} + \frac{\partial^2 V(x, y, z)}{\partial y^2} + \frac{\partial^2 V(x, y, z)}{\partial z^2} = 0. \quad (13.7)$$

The diffusion or heat flow equation: Suppose one has a function $u$ which describes the temperature at a given location $(x, y, z)$. This function will change over time as heat spreads throughout space. The heat equation is used to determine the change in the function $u$ over time. In a special case of heat flow in an isotropic and homogeneous medium in a 3-dimensional space, the heat equation is given by

$$\frac{\partial u}{\partial t} - \alpha \left[ \frac{\partial^2 u(x, y, z, t)}{\partial x^2} + \frac{\partial^2 u(x, y, z, t)}{\partial y^2} + \frac{\partial^2 V(x, y, z, t)}{\partial z^2} \right] = 0. \quad (13.8)$$

where $\alpha$ is the thermal diffusivity, a material-specific quantity depending on the thermal conductivity, the mass density, and the specific heat capacity

Wave equation: Consider a string with linear mass density $\lambda$ under tension force $T$ displaced from the equilibrium as shown in the figure below:

![Wave equation diagram](image)

If wave propagating on this string is described by the wave equation

$$\frac{\partial^2 f}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 f}{\partial t^2} \quad (13.9)$$

where

$$v = \sqrt{\frac{T}{\lambda}}. \quad (13.10)$$

is the speed of the wave.
13.2. LAPLACE’S EQUATION IN CARTESIAN COORDINATES

**Helmholtz Equation:** For three dimensional case the wave equation has the form

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \frac{1}{v^2} \frac{\partial^2 f}{\partial t^2} \quad (13.11)$$

The space dependent part that we would obtain using separation of variables has the form

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} + k^2 f = 0 \quad (13.12)$$

this is called Helmholtz equation.

**Schrödinger Equation:** in quantum mechanics the state of a particle is represented by the wave function $\psi(x, y, z, t)$ which satisfied the time dependent Schrödinger equation

$$V(x, y, z, t) \psi(x, y, z, t) - \frac{\hbar^2}{2m} \nabla^2 \psi(x, y, z, t) = i\hbar \frac{\partial}{\partial t} \psi(x, y, z, t) \quad (13.13)$$

We will study a technique commonly used to solve the PDEs discussed above. This technique is called the separation of variables technique. We will learn this technique by solving problems.

### 13.2 Laplace’s equation in Cartesian coordinates

Before we begin to solve Laplace’s equation in Cartesian coordinates, it is important to revise how we determine the solution for second order HLDE with constant coefficients.

**Example 13.1** Solve the differential equation

$$\frac{d^2 f(x)}{dx^2} = \pm k^2 f(x). \quad (13.14)$$

**Solution:** Assuming a solution of the form

$$f(x) = Ae^{\lambda x} \quad (13.15)$$

we find the indicial equation

$$\lambda^2 \mp k^2 = 0. \quad (13.16)$$

For the plus case the solutions are

$$\lambda_1 = k, \lambda_2 = -k \quad (13.17)$$

and the general solution is given by

$$f(x) = A_1 e^{-kx} + A_2 e^{kx}. \quad (13.18)$$
For the minus case the solution to the indicial equation are
\[ \lambda_1 = ik, \lambda_2 = -ik \] (13.19)
and the general solution becomes
\[ f(x) = B_1 \cos(kx) + B_2 \sin(kx). \] (13.20)

Laplace’s equation in Cartesian coordinates is given by
\[
\frac{\partial^2 V(x, y, z)}{\partial x^2} + \frac{\partial^2 V(x, y, z)}{\partial y^2} + \frac{\partial^2 V(x, y, z)}{\partial z^2} = 0.
\]

One can express \( V(x, y, z) \) as a product of three independent functions \( A(x), B(y), \) and \( C(z) \)
\[ V(x, y, z) = X(x)Y(y)Z(z). \] (13.21)

Substituting this expression into Laplace’s equation, we have
\[
Y(z) \frac{d^2 X(x)}{dx^2} + X(x) \frac{d^2 Y(y)}{dy^2} + X(x)Y(y) \frac{d^2 Z(z)}{dz^2} = 0.
\] (13.22)

and dividing by \( X(x)Y(y)Z(z) \), we find
\[
\frac{1}{X(x)} \frac{d^2 X(x)}{dx^2} + \frac{1}{Y(y)} \frac{d^2 Y(y)}{dy^2} + \frac{1}{Z(z)} \frac{d^2 Z(z)}{dz^2} = 0. \] (13.23)

This equation consists of three independent terms and the sum of these terms must be zero. This is possible if and only if these terms are constants. Therefore, we can write
\[
\frac{1}{X(x)} \frac{d^2 X(x)}{dx^2} = -k_x^2, \quad \frac{1}{Y(y)} \frac{d^2 Y(y)}{dy^2} = -k_y^2, \quad \frac{1}{Z(z)} \frac{d^2 Z(z)}{dz^2} = k_z^2,
\] (13.24)
so that
\[ k_x^2 + k_y^2 - k_z^2 = 0. \] (13.25)

These equations are 2-nd order ODE and can be solved using any of the techniques we have studied so far. Considering the \( x \) dependent part, we have
\[
\frac{1}{X(x)} \frac{d^2 X(x)}{dx^2} = -k_x^2 \Rightarrow \frac{d^2 X(x)}{dx^2} + k_x^2 X(x) = 0.
\] (13.26)

Similarly, for the \( y \) dependent part, we find
\[
\frac{d^2 Y(y)}{dy^2} + k_y^2 Y(y) = 0.
\] (13.27)
13.2. LAPLACE’S EQUATION IN CARTESIAN COORDINATES 363

Assuming \( k_x \) and \( k_y \) are real, the solutions to the above ODEs, using the results in Ex. 23, may be expressed in any one of the following form

\[
X_{k_x} (x) = \begin{cases} 
A_{k_x} e^{ik_xx} + B_{k_x} e^{-ik_xx}, \\
C_{k_x} \cos (k_xx) + D_{k_x} \sin (k_xx), \\
E_{k_x} \cos (k_xx - \gamma_{k_x}), \\
F_{k_x} \sin (k_xx + \beta_{k_x}), 
\end{cases}
\]

\[
Y_{k_y} (y) = \begin{cases} 
A_{k_y} e^{ik_yy} + B_{k_y} e^{-ik_yy}, \\
C_{k_y} \cos (k_yy) + D_{k_y} \sin (k_yy), \\
E_{k_y} \cos (k_yy - \gamma_{k_y}), \\
F_{k_y} \sin (k_yy + \beta_{k_y}), 
\end{cases}
\]  

(13.28)

For the \( z \) dependent part

\[
\frac{d^2 Z (z)}{dz^2} - k_z^2 Z (z) = 0, 
\]

(13.29)

assuming \( k_z \) is real, the solution can be expressed as

\[
Z_{k_z} (z) = \begin{cases} 
A_{k_z} e^{k_zz} + B_{k_z} e^{-k_zz}, \\
C_{k_z} \cosh (k_zz) + D_{k_z} \sinh (k_zz), \\
E_{k_z} \cosh (k_zz + \gamma_{k_z}), \\
F_{k_z} \sinh (k_zz + \beta_{k_z}), 
\end{cases}
\]  

(13.30)

Note: It is important to note that our choice of the constant in Eq. (13.25) can switch the solutions we found in Eqs. (13.27) and (13.30). For example, if we had chosen

\[
\frac{1}{X(x)} \frac{d^2 X(x)}{dx^2} = k_x^2, \quad \frac{1}{Y(y)} \frac{d^2 Y(y)}{dy^2} = k_y^2, \quad \frac{1}{Z(z)} \frac{d^2 Z(z)}{dz^2} = -k_z^2,
\]

\[
\Rightarrow \frac{d^2 X(x)}{dx^2} - k_x^2 X(x) = 0, \quad \frac{d^2 Y(y)}{dy^2} - k_y^2 Y(y) = 0, \quad \frac{d^2 Z(z)}{dz^2} + k_z^2 Z(z) = 0,
\]

(13.31)

where

\[
k_x^2 - k_y^2 - k_z^2 = 0. 
\]

(13.32)

The solutions would have have been

\[
X_{k_x'} (x) = \begin{cases} 
A_{k_x'} e^{ik_x'x} + B_{k_x'} e^{-ik_x'x}, \\
C_{k_x'} \cosh (k_x'x) + D_{k_x'} \sinh (k_x'x), \\
E_{k_x'} \cosh (k_x'x - \gamma_{k_x'}), \\
F_{k_x'} \sinh (k_x'x + \beta_{k_x'}), 
\end{cases}
\]

\[
Y_{k_y'} (y) = \begin{cases} 
A_{k_y'} e^{ik_y'y} + B_{k_y'} e^{-ik_y'y}, \\
C_{k_y'} \cosh (k_y'y) + D_{k_y'} \sinh (k_y'y), \\
E_{k_y'} \cosh (k_y'y - \gamma_{k_y'}), \\
F_{k_y'} \sinh (k_y'y + \beta_{k_y'}), 
\end{cases}
\]  

(13.33)
and

\[
Z_{k'_z} (z) = \left\{ \begin{array}{l}
A_{k'_z} e^{i k'_z z} + B_{k'_z} e^{-i k'_z z}, \\
C_{k'_z} \cos (k'_z z) + D_{k'_z} \sin (k'_z z), \\
E_{k'_z} \cos (k'_z z + \gamma_{k'_z}), \\
F_{k'_z} \sin (k'_z z + \beta_{k'_z}), \\
\end{array} \right.
\]

(13.34)

Although the solutions in Eqs. (13.33) and (13.34) appear to be different from Eqs. (13.28) and (13.30), one can find one from the other. This can be shown using

\[
k^2_z = -k^2_y, -k_y^2 = k^2_y - k^2_x = k^2_x,
\]

\Rightarrow k'_z = i k_z, k'_y = i k_y, k'_x = i k_x,
\]

(13.35)

and noting that

\[
\cosh (i k_x x) = \frac{e^{i k x x} + e^{-i k x x}}{2} = \cos (k_x x),
\]

\[
\sinh (i k_x x) = i \frac{e^{i k x x} - e^{-i k x x}}{2i} = i \sin (k_x x),
\]

which leads to

\[
X_{k'_z} (x) = \left\{ \begin{array}{l}
A_{k'_z} e^{i k_x x} + B_{k'_z} e^{-i k_x x}, \\
C_{k'_z} \cosh (k'_z x) + D_{k'_z} \sinh (k'_z x), \\
E_{k'_z} \cosh (k'_z x - \gamma_{k'_z}), \\
F_{k'_z} \sinh (k'_z x + \beta_{k'_z}), \\
\end{array} \right. = X_{k_z} (x),
\]

(13.36)

and

\[
Y_{k'_y} (y) = \left\{ \begin{array}{l}
A_{k'_y} e^{i k_y y} + B_{k'_y} e^{-i k_y y}, \\
C_{k'_y} \cosh (k'_y y) + D_{k'_y} \sinh (k'_y y), \\
E_{k'_y} \cosh (k'_y y - \gamma_{k'_y}), \\
F_{k'_y} \sinh (k'_y y + \beta_{k'_y}), \\
\end{array} \right. = Y_{k_y} (y),
\]

(13.37)

and

\[
Z_{k'_z} (z) = \left\{ \begin{array}{l}
A_{k'_z} e^{i k'_z z} + B_{k'_z} e^{-i k'_z z}, \\
C_{k'_z} \cos (k'_z z) + D_{k'_z} \sin (k'_z z), \\
E_{k'_z} \cos (k'_z z + \gamma_{k'_z}), \\
F_{k'_z} \sin (k'_z z + \beta_{k'_z}), \\
\end{array} \right. = Z_{k_z} (z),
\]

(13.38)

for

\[
k^2_z = -k^2_y, -k^2_y = k^2_y - k^2_x = k^2_x,
\]

(13.39)

The general solution to Laplace’s equation in Cartesian coordinates is given by

\[
V (x, y, z) = \sum_{k_x} \sum_{k_y} \sum_{k_z} X_{k_x} (x) Y_{k_y} (y) Z_{k_z} (z)
\]

(13.39)
13.2. LAPLACE'S EQUATION IN CARTESIAN COORDINATES

Note that the constants $k_x, k_y, k_z$ along with the other constants determined from the boundary condition set for the given problem and the condition

$$k_x^2 + k_y^2 - k_z^2 = 0.$$ \hfill (13.40)

We shall see how these constants are determined in the next example.

**Example 13.2** Consider an infinitely long, rectangular waveguide of width, $w$, and height $h$ with the top surface held at the constant potential, $V(y = h) = V_0$, and with all other sides grounded. A vacuum pump has been connected to the waveguide to remove most of the air within it. Find the electrostatic potential, $V(x, y, z)$, everywhere within the waveguide. Inside the the wave guide the electric potential satisfied Laplace’s equation

$$\frac{\partial^2 V(x, y, z)}{\partial x^2} + \frac{\partial^2 V(x, y, z)}{\partial y^2} + \frac{\partial^2 V(x, y, z)}{\partial z^2} = 0.$$ \hfill (13.41)

Figure 13.1: A rectangular wave guide with width $w$ and height $h$. It is infinitely long along the $z$ direction.

**Solution:** We already determined solution to Laplace’s equation in Cartesian coordinates. To find the potential inside the waveguide, we need to state the given boundary conditions and impose it to the general solution to determine the constants. To this end, we note that the general solution is given by

$$V(x, y, z) = \sum_{k_x, k_y, k_z} X_{k_x}(x) Y_{k_y}(y) Z_{k_z}(z),$$ \hfill (13.42)

where

$$k_x^2 + k_y^2 - k_z^2 = 0.$$ \hfill (13.43)
$X_{k_x}(x), Y_{k_y}(y)$, and $Z_{k_z}(x)$ are given by Eqs. (13.28) and (13.30). The boundary conditions are

$$V(0, y, z) = V(w, y, z) = V(x, 0, z) = 0, V(x, h, z) = V_0. \quad (13.44)$$

And also the potential must be finite everywhere inside the waveguide. Imposing these boundary conditions

$$V(0, y, z) = 0 \Rightarrow \sum_{k_x, k_y, k_z} X_{k_x}(0) Y_{k_y}(y) Z_{k_z}(z) = 0 \Rightarrow X_{k_x}(0) = 0,$$

$$V(w, y, z) = 0 \Rightarrow \sum_{k_x, k_y, k_z} X_{k_x}(w) Y_{k_y}(y) Z_{k_z}(z) = 0 \Rightarrow X_{k_x}(w) = 0,$$

$$V(x, 0, z) = 0 \Rightarrow \sum_{k_x, k_y, k_z} X_{k_x}(x) Y_{k_y}(0) Z_{k_z}(z) = 0 \Rightarrow Y_{k_y}(0) = 0 \quad (13.45)$$

Using the solutions in Eqs. (13.28), we find

$$X_{k_x}(0) = 0 \Rightarrow [C_{k_x} \cos (k_x x) + D_{k_x} \sin (k_x x)]_{x=0} = 0 \Rightarrow C_{k_x} = 0,$$

$$X_{k_x}(w) = 0 \Rightarrow [C_{k_x} \cos (k_x x) + D_{k_x} \sin (k_x x)]_{x=W} = 0 \Rightarrow C_{k_x} + D_{k_x} \sin (k_x w) = 0 \Rightarrow D_{k_x} \sin (k_x w) = 0 \Rightarrow k_x = \frac{n\pi}{w}, \text{ where } n = 0, 1, 2, 3...,,$$

$$Y_{k_y}(0) = 0 \Rightarrow [C_{k_y} \cos (k_y y) + D_{k_y} \sin (k_y y)]_{y=0} = 0 \Rightarrow C_{k_y} = 0. \quad (13.46)$$

Since the waveguide is infinitely long along the $z$-direction and the potential must be finite, we must have

$$\lim_{z \to \pm \infty} \left( \sum_{k_x, k_y, k_z} X_{k_x}(x) Y_{k_y}(y) Z_{k_z}(z) \right) \neq \pm \infty$$

$$\Rightarrow \lim_{z \to \pm \infty} Z_{k_z}(z) \neq \pm \infty \quad (13.47)$$

and using the result in Eq. (13.30), one finds

$$\lim_{z \to \pm \infty} \left[ A_{k_z} e^{k_z z} + B_{k_z} e^{-k_z z} \right] \neq \pm \infty \Rightarrow k_z = 0 \Rightarrow$$

$$Z_{k_z}(z) = \text{Constant} = A \quad (13.48)$$

Substituting $k_x = \frac{n\pi}{w}$ and $k_z = 0$ into

$$k_x^2 + k_y^2 - k_z^2 = 0, \quad (13.49)$$

we find

$$k_y = \pm i \frac{n\pi}{w}. \quad (13.50)$$
13.2. LAPLACE’S EQUATION IN CARTESIAN COORDINATES

At this stage we note

\[ X_k(x) = D_k \sin(kx) = D_n \sin\left(\frac{n\pi}{w}x\right), \]
\[ Y_k(y) = D_k \sin(ky) = D_k \sin\left(\pm i\frac{n\pi}{w}y\right) = \pm iD_k \sinh\left(\frac{n\pi}{w}y\right), \]
\[ Z_k(z) = A, \quad (13.51) \]

where we used the relation

\[ \sin(ix) = e^{i\pi x} - e^{-i\pi x} = -\frac{e^x - e^{-x}}{2i} = i\frac{e^x - e^{-x}}{2}, \]
\[ \Rightarrow \sin(ix) = i \sinh(x). \quad (13.52) \]

Then the general solution can be expressed as

\[ V(x, y, z) = \sum_{n=1}^{\infty} H_n \sin\left(\frac{n\pi}{w}x\right) \sinh\left(\frac{n\pi}{w}y\right). \quad (13.53) \]

Now applying the last boundary condition, \( V(x, h, z) = V_0 \), we have

\[ V_{k_n}(x, y = h, z) = V_0 \Rightarrow \sum_{n=1}^{\infty} H_n \sin\left(\frac{n\pi}{w}x\right) \sinh\left(\frac{n\pi}{w}h\right) = V_0 \quad (13.54) \]

where \( n = 1, 2, \ldots \) Multiplying both sides by \( \sin\left(\frac{n\pi}{w}x\right) \) and integrating with respect to \( x \) over the width of the waveguide, we have

\[ \sum_{n=1}^{\infty} H_n \sinh\left(\frac{n\pi}{w}h\right) \int_0^W \sin\left(\frac{n\pi}{w}x\right) \sin\left(\frac{m\pi}{w}x\right) dx = V_0 \int_0^W \sin\left(\frac{m\pi}{w}x\right) dx, \quad (13.55) \]

so that noting that

\[ \int_0^W \sin\left(\frac{n\pi}{w}x\right) \sin\left(\frac{m\pi}{w}x\right) dx = \frac{w}{\pi} \int_0^\pi \sin\left(\frac{n\pi}{w}x\right) \sin\left(\frac{m\pi}{w}x\right) d\left(\frac{x}{w}\right) \]
\[ = \frac{w}{\pi} \int_0^\pi \sin(nu) \sin(mu) du = \frac{w}{\pi} \frac{\pi}{2} \delta_{mn} = \frac{w}{2} \delta_{mn} \quad (13.56) \]

and

\[ \int_0^W \sin\left(\frac{m\pi}{w}x\right) dx = -\frac{w}{m\pi} \cos\left(\frac{m\pi}{w}x\right)|_0^W = \frac{w}{m\pi} [1 - (-1)^m] \quad (13.57) \]
we find
\[
\sum_{n=1}^{\infty} H_n \sinh \left( \frac{n\pi h}{w} \right) \frac{w}{2} \delta_{mn} = V_0 \frac{w}{m\pi} \left[ 1 - (-1)^m \right]
\]
\(\Rightarrow\) \(H_m \sinh \left( \frac{m\pi h}{w} \right) = V_0 \frac{2}{m\pi} \left[ 1 - (-1)^m \right]\)
\(\Rightarrow\) \(H_m = \left\{ \begin{array}{ll}
0, & m = \text{even} \\
\frac{4V_0}{m\pi \sinh(m\pi h/w)}, & m = \text{odd}
\end{array} \right. \quad (13.58)\)

Then the potential becomes
\[
V(x, y, z) = \frac{4V_0}{\pi} \sum_{n=0}^{\infty} \sin \left[ \frac{n\pi}{2} (2n + 1) \pi \right] \sinh \left[ \frac{w}{w} (2n + 1) \pi \right] \frac{1}{(2n + 1) \sinh \left[ \frac{n\pi}{w} (2n + 1) \pi \right]}. \quad (13.59)
\]

### 13.3 Laplace’s equation in spherical coordinates

We have seen an example in the previous lecture illustrating how we solve Laplace’s equation in Cartesian coordinates
\[
\nabla^2 V(x, y, z) = \frac{\partial^2 V(x, y, z)}{\partial x^2} + \frac{\partial^2 V(x, y, z)}{\partial y^2} + \frac{\partial^2 V(x, y, z)}{\partial z^2} = 0 \quad (13.60)
\]

In spherical coordinates, the Laplacian is given by
\[
\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left( \sin(\theta) \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2(\theta)} \frac{\partial^2}{\partial \phi^2} = 0.
\]

\(\Rightarrow\) \(\nabla^2 V = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial V}{\partial r} \right) + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left( \sin(\theta) \frac{\partial V}{\partial \theta} \right) + \frac{1}{r^2 \sin^2(\theta)} \frac{\partial^2 V}{\partial \phi^2} = 0 \quad (13.61)\)

Introducing separation of variables defined by
\[
V(r, \theta, \phi) = R(r) \Theta(\theta) \Phi(\phi) \quad (13.62)
\]

Laplace’s equation in spherical coordinates becomes
\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left[ r^2 \frac{\partial}{\partial r} \left( R(r) \Theta(\theta) \Phi(\phi) \right) \right] + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left[ \sin(\theta) \frac{\partial}{\partial \theta} \left( R(r) \Theta(\theta) \Phi(\phi) \right) \right] + \frac{1}{r^2 \sin^2(\theta)} \frac{\partial^2}{\partial \phi^2} \left( R(r) \Theta(\theta) \Phi(\phi) \right) = 0. \quad (13.63)
\]

\[
\frac{\Theta(\theta) \Phi(\phi)}{r^2} \frac{d}{dr} \left( r^2 \frac{dR(r)}{dr} \right) + R(r) \Phi(\phi) \frac{d}{d\theta} \left( \sin(\theta) \frac{d\Theta(\theta)}{d\theta} \right) + \frac{R(r) \Theta(\theta) \Phi(\phi)}{r^2 \sin^2(\theta)} = 0. \quad (13.64)
\]
Multiplying this equation by
\[ \frac{r^2}{R(r) \Theta(\theta) \Phi(\varphi)} \]
we find
\[ \frac{1}{R(r)} \frac{d}{dr} \left( r^2 \frac{dR(r)}{dr} \right) + \frac{1}{\sin(\theta) \Theta(\theta)} \frac{d}{d\theta} \left( \sin(\theta) \frac{d\Theta(\theta)}{d\theta} \right) \]
\[ + \frac{1}{\sin^2(\theta) \Phi(\varphi)} \frac{d^2 \Phi(\varphi)}{d\varphi^2} = 0. \] (13.65)

This can be put in the form
\[ F_1(r) + F_2(\theta, \varphi) = 0, \] (13.66)
where
\[ F_1(r) = \frac{1}{R(r)} \frac{d}{dr} \left( r^2 \frac{dR(r)}{dr} \right), \]
\[ F_2(\theta, \varphi) = \frac{1}{\sin(\theta) \Theta(\theta)} \frac{d}{d\theta} \left( \sin(\theta) \frac{d\Theta(\theta)}{d\theta} \right) + \frac{1}{\sin^2(\theta) \Phi(\varphi)} \frac{d^2 \Phi(\varphi)}{d\varphi^2} = 0. \] (13.67)
We note that $F_1 (r)$ depends on the variable, $r$ and $F_2 (\theta, \varphi)$ depends on the variables $\theta$ and $\varphi$. But the three variables $r$, $\theta$, and $\varphi$ are independent. Therefore, each of these function must be a constant,

$$F_1 (r) = k^2, F_2 (\theta, \varphi) = -k^2. \quad (13.68)$$

such that

$$F_1 (r) + F_2 (\theta, \varphi) = 0. \quad (13.69)$$

There follows that

$$\frac{1}{R(r)} \frac{d}{dr} \left( r^2 \frac{dR(r)}{dr} \right) = k^2 \Rightarrow \frac{d}{dr} \left( r^2 \frac{dR(r)}{dr} \right) - k^2 R(r) = 0$$

and

$$\frac{1}{\sin (\theta) \Theta(\theta)} \frac{d}{d\theta} \left( \sin (\theta) \frac{d\Theta(\theta)}{d\theta} \right) + \frac{1}{\sin^2 \theta \Phi(\varphi)} \frac{d^2 \Phi(\varphi)}{d\varphi^2} = -k^2$$

$$\Rightarrow \frac{\sin (\theta)}{\Theta(\theta)} \frac{d}{d\theta} \left( \sin (\theta) \frac{d\Theta(\theta)}{d\theta} \right) + \frac{1}{\Phi(\varphi)} \frac{d^2 \Phi(\varphi)}{d\varphi^2} = -k^2 \sin^2 \theta$$

$$\Rightarrow \sin (\theta) \frac{d}{d\theta} \left( \sin (\theta) \frac{d\Theta(\theta)}{d\theta} \right) + k^2 \sin^2 \theta + \frac{1}{\Phi(\varphi)} \frac{d^2 \Phi(\varphi)}{d\varphi^2} = 0. \quad (13.70)$$

We note that in the above expression the first two terms depend on $\theta$ and the third term depend on $\varphi$. Therefore, we must have

$$\frac{1}{\Phi(\varphi)} \frac{d^2 \Phi(\varphi)}{d\varphi^2} = -m^2, \Rightarrow \frac{d^2 \Phi(\varphi)}{d\varphi^2} + m^2 \Phi(\varphi) = 0,$$

$$\sin (\theta) \frac{d}{d\theta} \left( \sin (\theta) \frac{d\Theta(\theta)}{d\theta} \right) + k^2 \sin^2 \theta = m^2,$$

$$\Rightarrow \sin (\theta) \frac{d}{d\theta} \left( \sin (\theta) \frac{d\Theta(\theta)}{d\theta} \right) + \sin^2 \theta \left[ k^2 - \frac{m^2}{\sin^2 \theta} \right] \Theta(\theta) = 0. \quad (13.71)$$

The three equation resulting from Laplace’s equation in spherical coordinates are

$$\frac{d^2 \Phi(\varphi)}{d\varphi^2} + m^2 \Phi(\varphi) = 0,$$

$$\sin (\theta) \frac{d}{d\theta} \left( \sin (\theta) \frac{d\Theta(\theta)}{d\theta} \right) + \sin^2 \theta \left[ k^2 - \frac{m^2}{\sin^2 \theta} \right] \Theta(\theta) = 0,$$

$$r^2 \frac{d^2 R(r)}{dr^2} + 2r \frac{dR(r)}{dr} - k^2 R(r) = 0.$$

We now proceed to find the solutions for these equations.

(a) The function $\Phi(\varphi)$: we can easily find the solution to the differential equation

$$\frac{d^2 \Phi(\varphi)}{d\varphi^2} + m^2 \Phi(\varphi) = 0,$$  

$$r^2 \frac{d^2 R(r)}{dr^2} + 2r \frac{dR(r)}{dr} - k^2 R(r) = 0.$$

(13.72)
to be
\[ A_m \cos (m \varphi) + B_m \sin (m \varphi). \]

(b) The function \( \Theta (\theta) \): to determine the solution to the differential equation

\[
\sin (\theta) \frac{d}{d\theta} \left( \sin (\theta) \frac{d \Theta (\theta)}{d\theta} \right) + \sin^2 \theta \left[ k^2 - \frac{m^2}{\sin^2 \theta} \right] \Theta (\theta) = 0. \tag{13.73}
\]

we introducing the transformation of variables defined by

\[
x = \cos \theta \Rightarrow dx = -\sin (\theta) d\theta, \sin (\theta) = \sqrt{1 - x^2}
\]

\[
\Rightarrow d\theta = -\frac{dx}{\sin (\theta)} \Rightarrow \frac{d\theta}{\sin (\theta)} = -\frac{dx}{\sin^2 \theta} = -\frac{dx}{1 - x^2} \tag{13.74}
\]

so that one can rewrite Eq. (13.73) as

\[
-(1 - x^2) \frac{d}{dx} \left( - (1 - x^2) \frac{d \Theta (x)}{dx} \right) + (1 - x^2) \left[ k^2 - \frac{m^2}{1 - x^2} \right] \Theta (x) = 0
\]

\[
\Rightarrow (1 - x^2) \frac{d^2 \Theta (x)}{dx^2} - 2x \frac{d \Theta (x)}{dx} + \left[ k^2 - \frac{m^2}{1 - x^2} \right] \Theta (x) = 0 \tag{13.75}
\]

We recall the associated Legendre differential equation

\[
(1 - x^2) y'' - 2xy' + \left[ l (l + 1) - \frac{m^2}{1 - x^2} \right] y = 0 \tag{13.76}
\]

the solution of which we found to be the associate Legendre polynomials

\[
y (x) = P^m_l (x). \tag{13.77}
\]

Therefore, comparing Eqs. (13.75) and (13.76) one finds,

\[
k^2 = l (l + 1), \Theta_{lm} (\theta) = P^m_l (\cos \theta). \tag{13.78}
\]

(c) The function \( R (r) \): Next we proceed to find the solution for the differential equation for the \( r \) dependent part

\[
r^2 \frac{d^2 R (r)}{dr^2} + 2r \frac{dR (r)}{dr} - k^2 R (r) = 0.
\]

To this end using the result above

\[
k^2 = l (l + 1),
\]

one can write

\[
\frac{d}{dr} \left( r^2 \frac{dR (r)}{dr} \right) - l (l + 1) R (r) = 0
\]

\[
\Rightarrow r^2 \frac{d^2 R (r)}{dr^2} + 2r \frac{dR (r)}{dr} - l (l + 1) R (r) = 0
\]

\[
\Rightarrow \frac{d^2 R (r)}{dr^2} + \frac{2}{r} \frac{dR (r)}{dr} - \frac{l (l + 1)}{r^2} R (r) = 0. \tag{13.79}
\]
This equation has singularity at $r = 0$. In such cases we know that the solution is determined by using the Frobenius method. In this method we generally begin by assuming the solution can be expressed as power series that can be written as

$$R(r) = \sum_{n=0}^{\infty} a_n r^{n+s}$$

and determine both the expansion coefficients $a_n$ and the constant $s$. Substituting Eq. (e12.567a) and

$$\frac{dR(r)}{dr} = \sum_{n=0}^{\infty} a_n (n+s) x^{n+s-1}, \quad \frac{d^2R(r)}{dr^2} = \sum_{n=0}^{\infty} a_n (n+s) (n+s-1) x^{n+s-2}.$$  

into Eq. (13.79), we find

$$\sum_{n=0}^{\infty} a_n (n+s) (n+s-1) x^{n+s-2} + 2 \sum_{n=0}^{\infty} a_n (n+s) x^{n+s-2} - l (l+1) \sum_{n=0}^{\infty} a_n x^{n+s-2} = 0$$

$$\Rightarrow \sum_{n=0}^{\infty} a_n [(n+s) (n+s-1) + 2 (n+s) - l (l+1)] x^{n+s-2} = 0$$

$$\Rightarrow \sum_{n=0}^{\infty} a_n [(n+s) (n+s+1) - l (l+1)] x^{n+s-2} = 0.$$  

(13.81)

There follows that

$$(n+s) (n+s+1) - l (l+1) = 0 \Rightarrow s^2 + (2n+1)s + n(n+1) - l(l+1) = 0$$

which results in

$$s_1 = -(n+l+1), \quad s_2 = l - n$$  

(13.82)

Noting that for $s_1 = -[n+ (l+1)]$

$$\sum_{n=0}^{\infty} a_n r^{n+s_1} = \left( \sum_{n=0}^{\infty} a_n \right) \frac{r^l}{r^{l+1}} = A_l \frac{r^l}{r^{l+1}}$$

(13.84)

and for $s_2 = l - n$

$$\sum_{n=0}^{\infty} a_n r^{n+s_2} = \left( \sum_{n=0}^{\infty} a_n \right) r^l = A_l r^l.$$  

(13.85)

one can easily see that the general solution to Eq. (13.79) is given by

$$R_l(r) = \sum_{n=0}^{\infty} a_n r^{n+s} = B_l r^l + \frac{C_l}{r^{l+1}}.$$  

(13.86)
Therefore the solution to Laplace’s equation in spherical coordinates as

\[ V(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} R_l(r) \Theta_{lm}(\theta) \Phi_{m}(\varphi) \quad (13.87) \]

or

\[ V(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left[ B_l r^l + \frac{C_l}{r^{l+1}} \right] P_l^m(\cos \theta) [A_m \cos(m\varphi) + B_m \sin(m\varphi)]. \quad (13.88) \]

This is the general solution for Laplace’s equation in spherical coordinates. Note that we set \( l = 0, 1, 2, \ldots \)
as we have already known from the solution to the Legendre differential equation. But how do we know that \( m = -l, l+1, \ldots -1, 0, 1, \ldots l-1 \)?

Let’s see if we can show that. To this end, we note that for the maximum value of \( x \), \( x = 1 \), since

\[ l(l+1) - \frac{m^2}{1-x^2} \]
diverges we must have

\[ m^2 = 0 \Rightarrow m = 0. \quad \text{(13.89)} \]

On the other hand for the minimum value of \( x \), \( x = 0 \), we have

\[ l(l+1) - \frac{m^2}{1-x^2} = l(l+1) - m^2 \quad \text{(13.90)} \]

which leads to

\[ l(l+1) - m^2 \geq 0 \Rightarrow m^2 \leq l(l+1) \]
\[ \Rightarrow -\sqrt{l(l+1)} \leq m \leq \sqrt{l(l+1)} \quad \text{(13.91)} \]

since \( m \) is an integer one finds

\[ -l \leq m \leq l. \quad \text{(13.92)} \]

**Example 13.3** All of space is initially filled with a uniform electric field of magnitude, \( E_0 \), pointing in the positive z-direction as shown in Fig. 13.3. A grounded conducting sphere of radius, \( a \), is introduced into the space with its center at the origin of coordinates. We wish to find the electrostatic potential at all points outside of the sphere.
CHAPTER 13. A PARTIAL DIFFERENTIAL EQUATION

Figure 13.3: A grounded metallic sphere in a uniform electric field.

Solution: Since there is no free charge outside the sphere, the electrical potential $V(\vec{r})$ satisfies the Laplace’s equation. Because of the spherical symmetry, it is easy to solve the Laplace’s equation in spherical coordinates,

\[
\nabla^2 V(\vec{r}) = 0
\]

\[
\Rightarrow \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} V(r,\theta,\varphi) \right) + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left( \sin(\theta) \frac{\partial}{\partial \theta} V(r,\theta,\varphi) \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} V(r,\theta,\varphi) = 0. \quad (13.93)
\]

While solving this partial differential equation I will outline the basic steps which we used in the previous example. It can be used as a general guideline to solve partial differential equations.

Identify and write down the boundary conditions for the given problem: For this problem we know that the sphere is grounded. Hence, the electric potential at any point on the surface of the sphere must be zero. That means

\[
V(r = a, \theta, \varphi) = 0. \quad (13.94)
\]

In addition we should expect that the presence of the conducting sphere far away from the sphere (infinity) is negligible. In other words at infinity the electric potential is just the electric potential of a constant electric
field pointing along the z-direction. Recalling that the electric potential \( V \) and electric field \( \vec{E} \) are related by

\[
\vec{E} = -\nabla V,
\]

(13.95)

for an electric field pointing along the positive z-direction, we have

\[
\vec{E} = E_0 \hat{z} = -\nabla V (x \to \infty, y \to \infty, z \to \infty) \\
\Rightarrow V (x \to \infty, y \to \infty, z \to \infty) = -E_0 z.
\]

(13.96)

We need to express this potential in spherical coordinates

\[
V (r, \theta, \varphi) = E_0 r \cos \theta.
\]

(13.97)

Apply the boundary conditions to the general solution: The general solution is given by

\[
V (r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{B_l}{l! \cdot d^{l+1}} P_l^m (\cos \theta) [A_m \cos (m\varphi) + B_m \sin (m\varphi)].
\]

(13.98)

The sphere is grounded

\[
V (r = a, \theta, \varphi) = 0,
\]

(13.99)

which means

\[
V (a, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left( B_l a^l + \frac{C_l}{a^{l+1}} \right) P_l^m (\cos \theta) \times (A_m \cos (m\varphi) + B_m \sin (m\varphi)) = 0 \Rightarrow B_l a^l + \frac{C_l}{a^{l+1}} = 0 \Rightarrow C_l = -B_l a^{2l+1}.
\]

(13.100)

Using this result we may want to rewrite the potential as

\[
V (r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B_l \left( r^l - \frac{a^{2l+1}}{r^{l+1}} \right) P_l^m (\cos \theta) \times (A_m \cos (m\varphi) + B_m \sin (m\varphi)).
\]

(13.101)

Now using the second boundary condition

\[
V (r \to \infty, \theta, \varphi) = -E_0 r \cos \theta
\]

(13.102)

we find

\[
\lim_{r \to \infty} V (r, \theta, \varphi) = \lim_{r \to \infty} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B_l r^l P_l^m (\cos \theta) (A_m \cos (m\varphi) + B_m \sin (m\varphi)) = -E_0 r \cos \theta
\]

(13.103)
Since there is no \( \varphi \) dependence on the right hand side, we must have
\[
m = 0, \quad (13.104)
\]
which leads to
\[
\lim_{r \to \infty} \sum_{l=0}^{\infty} A_0 B_l r^l P_l (\cos \theta) = -E_0 r \cos \theta = E_0 r P_1 (\cos \theta). \quad (13.105)
\]
Expanding the series and including the constant \( A_0 \) into \( B_l \), we have
\[
B_0 + B_1 r P_1 (\cos \theta) + \lim_{r \to \infty} \sum_{l=2}^{\infty} B_l r^l P_l (\cos \theta) = -E_0 r P_1 (\cos \theta). \quad (13.106)
\]
Comparing the right and left hand side of this equation, we find
\[
B_1 = -E_0, B_l = 0 \text{ for } l \neq 1. \quad (13.107)
\]
Substituting this result into the reduced expression for the potential
\[
V (r, \theta, \varphi) = \sum_{l=0}^{\infty} B_l \left( r^l - \frac{a^{2l+1}}{r^{l+1}} \right) P_l (\cos \theta), \quad (13.108)
\]
we find
\[
V (r, \theta, \varphi) = \left( \frac{a^3}{r^3} - 1 \right) E_0 r \cos \theta. \quad (13.109)
\]

### 13.4 Laplace’s equation in cylindrical coordinates

Consider the Laplace equation in cylindrical coordinates,
\[
\nabla^2 V = \frac{1}{s} \frac{\partial}{\partial s} \left( s \frac{\partial V}{\partial s} \right) + \frac{1}{s^2} \frac{\partial^2 V}{\partial \varphi^2} + \frac{\partial^2 V}{\partial z^2} = 0 \quad (13.110)
\]
Using separation of variables
\[
V (s, \varphi, z) = R (s) \phi (\varphi) Z (z)
\]
we have
\[
\phi (\varphi) Z (z) \frac{d}{ds} \left( s \frac{dR (s)}{ds} \right) + \frac{R (s) Z (z)}{s^2} \frac{d^2 \phi (\varphi)}{d\varphi^2} + R (s) \phi (\varphi) \frac{d^2 Z (z)}{dz^2} = 0 \quad (13.111)
\]
so that multiplying by \( s^2 / R (s) \phi (\varphi) Z (z) \), we find
\[
s^2 \left[ \frac{1}{sR (s)} \frac{d}{ds} \left( s \frac{dR (s)}{ds} \right) + \frac{1}{Z (z)} \frac{d^2 Z (z)}{dz^2} \right] + \frac{1}{\phi (\varphi)} \frac{d^2 \phi (\varphi)}{d\varphi^2} = 0, \quad (13.112)
\]
that leads to

\[ s^2 \left[ \frac{1}{sR(s)} \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) + \frac{1}{Z(z)} \frac{d^2Z(z)}{dz^2} \right] = p^2, \]

\[ \frac{1}{\phi(\varphi)} \frac{d^2\phi(\varphi)}{d\varphi^2} = -p^2. \tag{13.113} \]

We can write the differential equation

\[ s^2 \left[ \frac{1}{sR(s)} \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) + \frac{1}{Z(z)} \frac{d^2Z(z)}{dz^2} \right] = p^2 \tag{13.114} \]

as

\[ \frac{1}{sR(s)} \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) + \frac{1}{Z(z)} \frac{d^2Z(z)}{dz^2} = \frac{p^2}{s^2} \]

\[ \frac{1}{sR(s)} \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) - \frac{p^2}{s^2} + \frac{1}{Z(z)} \frac{d^2Z(z)}{dz^2} = 0 \tag{13.115} \]
so that
\[
\frac{1}{s R(s)} \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) - \frac{p^2}{s^2} = - \left( \frac{q}{r_0} \right)^2,
\]
\[
\frac{1}{Z(z)} \frac{d^2 Z(z)}{dz^2} = \left( \frac{q}{r_0} \right)^2, \quad (13.116)
\]
where \( q \) is a constant and \( r_0 \) is the radius of the cylinder. The differential equation
\[
\frac{1}{s R(s)} \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) - \frac{p^2}{s^2} = - \left( \frac{q}{r_0} \right)^2
\]
can be rewritten as
\[
s \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) - p^2 R(s) = - \left( \frac{q}{r_0} \right)^2 s^2 R(s)
\]
\[
\Rightarrow s \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) + \left[ \left( \frac{qs}{r_0} \right)^2 - p^2 \right] R(s) = 0
\]
\[
\Rightarrow s^2 \frac{d^2 R(s)}{ds^2} + s \frac{dR(s)}{ds} + \left[ \left( \frac{qs}{r_0} \right)^2 - p^2 \right] R(s) = 0, \quad (13.118)
\]
which we put in the form
\[
(qs)^2 \frac{d^2 R(s)}{d(qs)^2} + (qs) \frac{dR(s)}{d(qs)} + \left[ \left( \frac{qs}{r_0} \right)^2 - p^2 \right] R(s) = 0 \quad (13.119)
\]
Introducing the transformation of variable defined by, \( r = qs \), we have
\[
r^2 \frac{d^2 R(r)}{dr^2} + r \frac{dR(r)}{dr} + \left( \frac{r^2}{r_0^2} - p^2 \right) R(r) = 0 \quad (13.120)
\]
If the radius of the cylinder is, \( r_o \), then we can introduce a dimensionless variable,
\[
x = \frac{r}{r_o}, dr = r_o dx, R(r) = y(x) \quad (13.121)
\]
so that
\[
x^2 \frac{d^2 y(x)}{dx^2} + x \frac{dy(x)}{dx} + (x^2 - p^2) y(x) = 0. \quad (13.122)
\]
Since
\[
0 \leq r \leq r_o, \text{ and } x = \frac{r}{r_o} \quad (13.123)
\]
the solution to the differential equation
\[
x^2 \frac{d^2 y(x)}{dx^2} + x \frac{dy(x)}{dx} + (x^2 - p^2) y(x) = 0. \quad (13.124)
\]
exists only for $0 \leq x \leq 1$. This is the Bessel DE and its solutions are given by the Bessel functions
\begin{equation}
y (x) = a_p J_p (x) + b_p N_p (x). \tag{13.125}
\end{equation}
Noting that the solution for the $\varphi$ dependent of the differential equation
\begin{equation}
\frac{1}{\phi (\varphi)} \frac{d^2 \phi (\varphi)}{d \varphi^2} = -p^2 \tag{13.126}
\end{equation}
the solution is given by
\begin{equation}
\phi (\varphi) = A_p \cos (p \varphi) + B_p \sin (p \varphi) \tag{13.127}
\end{equation}
and for the $z$-dependence
\begin{equation}
\frac{1}{Z (z)} \frac{d^2 Z (z)}{d z^2} = \left( \frac{q}{r_0} \right)^2 \tag{13.128}
\end{equation}
the solution is given by
\begin{equation}
Z (z) = C_q e^{\frac{q}{r_0}z} + D_q e^{-\frac{q}{r_0}z}. \tag{13.129}
\end{equation}
The trigonometric functions must be single valued, this requires,
\begin{equation}
\cos (p \varphi) = \cos [p (\varphi + 2\pi)] , \sin (p \varphi) = \sin [p (\varphi + 2\pi)] \tag{13.130}
\end{equation}
As one can see from Fig. , this requires $p$ must be an integer, $p = n = 1, 2, 3, \ldots$. Therefore the general solution for the Laplace equation is given by
\begin{equation}
V (s, \varphi, z) = \sum_q \sum_{n=1} \left( C_q e^{\frac{q}{r_0}z} + D_q e^{-\frac{q}{r_0}z} \right) \left( A_n \cos (n \varphi) + B_n \sin (n \varphi) \right) J_n \left( \frac{qs}{r_o} \right). \tag{13.131}
\end{equation}
Note that we have taken into account that, when the constant $p$ is an integer, $p = n$, $J_n (x)$ and $J_{-n} (x)$ are not independent
\begin{equation}
J_{-n} (x) = (-1)^n J_n (x). \tag{13.131}
\end{equation}
Note that the Laplace’s equation
\begin{equation}
s^2 \left[ \frac{1}{sR (s)} \frac{d}{ds} \left( s \frac{dR (s)}{ds} \right) + \frac{1}{Z (z)} \frac{d^2 Z (z)}{d z^2} \right] + \frac{1}{\phi (\varphi)} \frac{d^2 \phi (\varphi)}{d \varphi^2} = 0, \tag{13.132}
\end{equation}
using
\begin{equation}
\frac{1}{Z (z)} \frac{d^2 Z (z)}{d z^2} = \left( \frac{q}{r_0} \right)^2 , \frac{1}{\phi (\varphi)} \frac{d^2 \phi (\varphi)}{d \varphi^2} = -p^2 \tag{13.132}
\end{equation}
we have
\begin{equation}
s^2 \left[ \frac{1}{sR (s)} \frac{d}{ds} \left( s \frac{dR (s)}{ds} \right) + \left( \frac{q}{r_0} \right)^2 \right] - p^2 = 0, \tag{13.133}
\end{equation}
Figure 13.5: The functions $\sin(p\varphi)$ (blue dashed line) and $\sin(p(\varphi + 2\pi))$ (red dotted line) when $p$ is an integer (in this case $p = 1$)

Figure 13.6: The functions $\sin(p\varphi)$ (blue dashed line) and $\sin(p(\varphi + 2\pi))$ (red dotted line) when $p$ is not an integer (in this case $p = 0.5$)
so that for \( p = 0 \) and \( q = 0 \), we find

\[
\frac{s}{R(s)} \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) = 0 \Rightarrow \frac{d}{ds} \left( s \frac{dR(s)}{ds} \right) = 0, \Rightarrow s \frac{dR(s)}{ds} = A \\
\Rightarrow dR(s) = A \frac{ds}{s} \Rightarrow R(s) = A \ln(s)
\]

Therefore the general solution to Laplace’s equation can be expressed as

\[
V(s, \varphi, z) = A_0 \ln(s) \\
+ \sum_{q>0} \sum_{n>0} \left( C_q e^{\frac{z}{r_0}} + D_q e^{-\frac{z}{r_0}} \right) \left( A_n \cos(n \varphi) + B_n \sin(n \varphi) \right) J_p \left( \frac{qs}{r_0} \right).
\]

(13.135)

**Example 12.4** A right, circular conducting cylindrical shell of radius \( r_0 \) and length \( L \) has its axis coincident with the \( z \)-axis and its ends at \( z = 0 \) and \( z = L \). All sides of the cylinder are grounded except for the face at \( z = 0 \), which is maintained at a position-dependent potential specified by

\[
V(s, \varphi, z = 0) = s \cos \varphi,
\]

(13.136)

(see Fig. 13.4). Find the electrostatic potential everywhere inside the cylinder.

**Solution:** We use the general solution for Laplace’s equation in cylindrical
CHAPTER 13. A PARTIAL DIFFERENTIAL EQUATION

coordinate, which is given by

\[ V(s, \varphi, z) = A \ln(s) + \sum_{q>0} \sum_{n>0} \left( C_q e^{\frac{\varphi}{r_0}} + D_q e^{-\frac{\varphi}{r_0}} \right) \left( A_n \cos(n\varphi) + B_n \sin(n\varphi) \right) J_n \left( \frac{qs}{r_o} \right). \]  

(13.137)

This potential must satisfy the boundary conditions

\[ V(s, \varphi, 0) = s \cos(\varphi), V(s, \varphi, L) = 0, \text{ and } V(r_0, \varphi, z) = 0. \]  

(13.138)

Applying the boundary condition, \( V(s, \varphi, 0) = 0 \), we have

\[ V(s, \varphi, 0) = A_0 \ln(s) + \sum_{q>0} \sum_{n>0} \left( C_q + D_q \right) \left( A_n \cos(n\varphi) + B_n \sin(n\varphi) \right) J_n \left( \frac{qs}{r_o} \right) = s \cos(\varphi). \]  

(13.139)

which leads to

\[ A_n = 0, \text{ for } n \neq 1, B_n = 0, \text{ for } n > 1 \Rightarrow J_n \left( \frac{qs}{r_o} \right) = J_1 \left( \frac{qs}{r_o} \right). \]  

(13.140)

Therefore, we can rewrite the potential as

\[ V(s, \varphi, z) = \sum_q \left( C_q e^{\frac{\varphi}{r_0}} + D_q e^{-\frac{\varphi}{r_0}} \right) \cos(\varphi) J_1 \left( \frac{qs}{r_o} \right), \]  

(13.141)

where we included \( A_1 \) into the constants, \( C_q \) and \( D_q \). Using the second boundary condition, \( V(s, \varphi, L) = 0 \),

\[ V(s, \varphi, L) = \sum_q \left( C_q e^{\frac{L}{r_0}} + D_q e^{-\frac{L}{r_0}} \right) \cos(\varphi) J_1 \left( \frac{qs}{r_o} \right) = 0. \]  

(13.142)

we find

\[ C_q e^{\frac{L}{r_0}} + D_q e^{-\frac{L}{r_0}} = 0 \Rightarrow D_q = -C_q e^{2\frac{L}{r_0}}. \]  

(13.143)

Then the simplified expression for the potential can be rewritten as

\[ V(s, \varphi, z) = \sum_q C_q \left( e^{\frac{\varphi}{r_0}} - e^{2\frac{L}{r_0}} e^{-\frac{\varphi}{r_0}} \right) \cos(\varphi) J_1 \left( \frac{qs}{r_o} \right). \]  

(13.144)

or

\[ V(s, \varphi, z) = \sum_q 2C_q e^{\frac{\varphi}{r_0}} \left( e^{-\frac{\varphi}{r_0}(L-z)} - e^{\frac{\varphi}{r_0}(L-z)} \right) \frac{1}{2} \cos(\varphi) J_p \left( \frac{qs}{r_o} \right) \]

\[ \Rightarrow V(s, \varphi, z) = \sum_q F_q \sinh \left[ \frac{q}{r_0} (L - z) \right] \cos(\varphi) J_1 \left( \frac{qs}{r_o} \right). \]  

(13.145)
13.4. LAPLACE’S EQUATION IN CYLINDRICAL COORDINATES

Now we apply the last boundary condition, \( V(s = r_0, \varphi, z) = 0 \),

\[
V(r_0, \varphi, z) = \sum_q F_q \sinh \left( \frac{q}{r_0} (L - z) \right) \cos(\varphi) J_1(q) = 0, \quad (13.146)
\]

which leads to

\[
J_1(q) = 0, \quad (13.147)
\]

which means

\[
q = \alpha_i, \text{ for } i = 1, 2, 3...
\]

are the zeros of the Bessel function. Thus the potential can be written as

\[
V(s, \varphi, z) = \sum_i F_i \sinh \left( \frac{\alpha_i}{r_0} (L - z) \right) \cos(\varphi) J_1 \left( \frac{\alpha_i s}{r_o} \right), \quad (13.149)
\]

To find \( F_i \), first we reapply the first boundary condition

\[
V(s, \varphi, z = 0) = s \cos \varphi
\]

so that one can write

\[
V(s, \varphi, z = 0) = \sum_i F_i \sinh \left( \frac{\alpha_i}{r_0} L \right) \cos(\varphi) J_1 \left( \frac{\alpha_i s}{r_o} \right) = s \cos \varphi
\]

\[
\Rightarrow \sum_i F_i \sinh \left( \frac{\alpha_i}{r_0} L \right) J_1 \left( \frac{\alpha_i s}{r_o} \right) = s. \quad (13.151)
\]

We then employ the orthogonality relation for the Bessel function

\[
\langle J_p(\alpha_i x) | x J_p(\alpha_j x) \rangle = \int_0^1 J_p(\alpha_i x) J_p(\alpha_j x) x dx = \frac{J_p'(\alpha_i)}{2} \delta_{ij} \delta_{ij}.
\]

Multiplying both sides by \( J_1 \left( \frac{\alpha_i s}{r_o} \right) \) and integrating over \( s \), we have

\[
\sum_i F_i \sinh \left( \frac{\alpha_i}{r_0} L \right) \int_0^{r_0} J_1 \left( \frac{\alpha_i s}{r_o} \right) J_1 \left( \frac{\alpha_j s}{r_o} \right) \frac{s}{r_o} d \left( \frac{s}{r_o} \right)
\]

\[
= \int_0^{r_0} s J_1 \left( \frac{\alpha_i s}{r_o} \right) \frac{s}{r_o} d \left( \frac{s}{r_o} \right). \quad (13.153)
\]

and replacing, \( \frac{s}{r_o} = x \), we find

\[
\sum_i F_i \sinh \left( \frac{\alpha_i}{r_o} L \right) \int_0^1 J_1(\alpha_i x) J_1(\alpha_j x) x d (x) = \int_0^1 r_0 x J_1(\alpha_j x) x d (x)
\]

\[
\Rightarrow \sum_i F_i \sinh \left( \frac{\alpha_i}{r_o} L \right) J_p'(\alpha_i) J_p'(\alpha_j) \delta_{ij} = \int_0^1 r_0 x J_1(\alpha_j x) x d x.
\]

\[
\Rightarrow F_j \sinh \left( \frac{\alpha_j}{r_o} L \right) \frac{J_p^2(\alpha_j)}{2} = \int_0^1 r_0 x^2 J_1(\alpha_j x) d x \Rightarrow F_j = \frac{r_0 \int_0^1 x^2 J_1(\alpha_j x) d x}{\sinh \left( \frac{\alpha_j}{r_o} L \right) \frac{J_p^2(\alpha_j)}{2}} \quad (13.154)
\]
Using \textit{Mathematica} one can show that
\[
\int_0^1 x^2 J_1(\alpha_j x) \, dx = \frac{J_2(\alpha_j)}{\alpha_j} \quad (13.155)
\]
and
\[
F_j = \frac{r_0 J_2(\alpha_j)}{\alpha_j \sinh \left( \frac{\alpha_j L}{r_0} \right) J_1^2(\alpha_j) / 2} = \frac{2r_0}{\alpha_j \sinh \left( \frac{\alpha_j L}{r_0} \right) J_2(\alpha_j)} \quad (13.156)
\]
where we used the relation
\[
J'_j(\alpha_j) = J_{P+1}(\alpha_j). \quad (13.157)
\]
Therefore, the potential inside the cylinder becomes
\[
V(s, \varphi, z) = \sum_{j=1}^{\infty} \frac{2r_0 \sinh \left( \frac{\alpha_j}{r_0} (L - z) \right) J_1 \left( \frac{\alpha_j}{r_0} L \right)}{\alpha_j \sinh \left( \frac{\alpha_j}{r_0} L \right) J_2(\alpha_j)} \cos(\varphi),
\]
where \(q_j\) are the zeroes of the Bessel function.

### 13.5 Poisson’s Equation

In the previous three lectures we have seen several examples from electrostatics to determine the electric potential of a given charge distribution in a region where there is no charge with specific boundary conditions. In such cases the electric potential satisfy Laplace’s equation. We have seen how to solve Laplace’s equation in Cartesian coordinates where it is given by
\[
\frac{\partial^2}{\partial x^2} V(x, y, z) + \frac{\partial^2}{\partial y^2} V(x, y, z) + \frac{\partial^2}{\partial z^2} V(x, y, z) = 0, \quad (13.158)
\]
in cylindrical coordinates
\[
\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial}{\partial r} V(r, \varphi, z) \right) + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} V(r, \varphi, z) + \frac{\partial^2}{\partial z^2} V(r, \varphi, z) = 0, \quad (13.159)
\]
and in spherical coordinates
\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} V(r, \theta, \varphi) \right) + \frac{1}{r^2 \sin(\theta)} \left[ \sin(\theta) \frac{\partial}{\partial \theta} \left( \sin(\theta) \frac{\partial}{\partial \theta} V(r, \theta, \varphi) \right) + \frac{\partial^2}{\partial \varphi^2} V(r, \theta, \varphi) \right] = 0. \quad (13.160)
\]
We now see how we determine the electric potential in a region where there is some volume charge distribution, \( \rho(\hat{r}) \). In such cases the electric potential, \( V(x, y, z) \), satisfies the Poisson’s Equation given by
\[
\nabla^2 V(x, y, z) = -\frac{\rho(\hat{r})}{\epsilon_0}, \quad (13.161)
\]
where $\varepsilon_0$ is the electrical permittivity of a free space. I will discuss the basic procedures for solving Poisson’s Equation and applying the boundary conditions using an example in spherical coordinates.

**Approach to Solving Poisson’s Equation:** when we studied how to solve inhomogenous second order linear differential equations, for example

$$\frac{d^2 y}{dx^2} + a \frac{dy}{dx} + f (x), \quad (13.162)$$

we have seen that the solution is sum of the homogenous $y_h$ and the particular $y_p$ solutions given by

$$y = y_h + y_p. \quad (13.163)$$

The Homogenous solution, $y_h$, satisfies the equation

$$\frac{d^2 y_h}{dx^2} + a \frac{dy_h}{dx} + y_h = 0 \quad (13.164)$$

and the particular solution, $y_p$, is determined using the different techniques we discussed. The same principle is applied in solving Poisson’s equation. For the Poisson’s equation

$$\nabla^2 V(x, y, z) = \frac{\rho(\vec{r})}{\varepsilon_0} \quad (13.165)$$

the homogenous solutions, $V_h$, is basically is the solution to the Laplace equation

$$\nabla^2 V_h(x, y, z) = 0 \quad (13.166)$$

which we already know how to determine the solution. Suppose we determine the particular solution, $V_p$, then the solutions is given by

$$V(x, y, z) = V_h(x, y, z) + V_p(x, y, z). \quad (13.167)$$

We can easily show that this solution satisfies the Poisson’s equation

$$\nabla^2 V(x, y, z) = \nabla^2 [V_h(x, y, z) + V_p(x, y, z)] = \nabla^2 V_h(x, y, z) + \nabla^2 V_p(x, y, z) \quad (13.168)$$

Since $V_h$ is the solution to Laplace’s equation, we have

$$\nabla^2 V_h = 0, \quad (13.169)$$

and $V_p$ is the particular solution

$$\nabla^2 V_p = -\frac{\rho(\vec{r})}{\varepsilon_0} \quad (13.170)$$

we can see that

$$\nabla^2 V(x, y, z) = -\frac{\rho(\vec{r})}{\varepsilon_0}. \quad (13.171)$$

Once we determine the homogenous and particular solutions we then write the general solution and apply the boundary conditions.
Example 12.5 Point Charge and Grounded Conducting Sphere Note: A point charge $Q$ is located a distance, $a$, from the center of a grounded, conducting sphere of radius $R$, where $R < a$. Find the electrostatic potential everywhere outside of the sphere. Assume that the center of the sphere is at the origin of coordinates, and that the point charge is on the positive $z$ axis.

![Figure 13.7: A point charge, $Q$, and a grounded conducting sphere of radius, $R$.](image)

Solution: We want to find the electric potential at some point outside the sphere $r > R$. Since there is a point charge along the $z$ axis at a distance $a$ (> $R$) which is in the region outside the sphere we obviously see that the charge density is not zero in this region and the Laplace’s equation is not valid in all regions outside the sphere. Therefore, to find the electric potential we need to find the solution to the Poisson’s Equation. From the nature of the problem it is better to use spherical coordinates

$$\nabla^2 V(r, \theta, \varphi) = -\frac{\rho(\hat{r})}{\epsilon_0}$$

$$\Rightarrow \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} V(r, \theta, \varphi) \right) + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left( \sin(\theta) \frac{\partial}{\partial \theta} V(r, \theta, \varphi) \right)$$

$$+ \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} V(r, \theta, \varphi) = -\frac{\rho(\hat{r})}{\epsilon_0}. \quad (13.172)$$
The Homogenous solution for this equation is the general solution to the Laplace’s equation in spherical coordinates
\[
\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial V_h(r, \theta, \varphi)}{\partial r} \right) + \frac{1}{r^2 \sin(\theta)} \frac{\partial}{\partial \theta} \left( \sin(\theta) \frac{\partial V_h(r, \theta, \varphi)}{\partial \theta} \right) \\
+ \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} V_h(r, \theta, \varphi) = 0
\] (13.173)
and we recall that the general solution is given by
\[
V_h(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left( B_l r^l + \frac{C_l}{r^{l+1}} \right) P_l^m(\cos \theta) \left( A_m \cos(m \varphi) + B_m \sin(m \varphi) \right).
\] (13.174)
We recall that for a point charge, \(Q\), located at a position, \(\vec{r}'\), the electric potential at a distance, \(\vec{r}\), is given by
\[
V_Q(r, \theta, \varphi) = \frac{1}{4\pi \varepsilon_0} \frac{Q}{|\vec{r} - \vec{r}'|}.
\] (13.175)
For
\[
\vec{r} = r \sin(\theta) \cos \varphi \hat{x} + r \sin(\theta) \sin \varphi \hat{y} + r \cos \theta \hat{z}
\]
and
\[
\vec{r}' = a \hat{z},
\]
we have
\[
|\vec{r} - \vec{r}'| = (r^2 + a^2 - 2ra \cos \theta)^{1/2}
\] (13.176)
so that
\[
V_Q(r, \theta, \varphi) = \frac{1}{4\pi \varepsilon_0} \frac{Q}{(r^2 + a^2 - 2ra \cos \theta)^{1/2}}.
\] (13.177)
To find the particular solution we note that the charge outside the sphere is zero except the point charge at a point on the \(z\) axis a distance \(a\) from the origin. Therefore, the particular solution for the Poisson’s equation
\[
\nabla^2 V(r, \theta, \varphi) = -\frac{\rho(r)}{\varepsilon_0}
\] (13.178)
is that of the electrical potential of this point charge
\[
V_p(r, \theta, \varphi) = V_Q(r, \theta, \varphi) = \frac{1}{4\pi \varepsilon_0} \frac{Q}{(r^2 + a^2 - 2ra \cos \theta)^{1/2}}.
\] (13.179)
The general solution to the Poisson’s equation can then be written as
\[
V(r, \theta, \varphi) = V_h(r, \theta, \varphi) + V_p(r, \theta, \varphi)
\]
\[
V(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left( B_l r^l + \frac{C_l}{r^{l+1}} \right) P_l^m(\cos \theta) \\
\times \left( A_m \cos(m \varphi) + B_m \sin(m \varphi) \right) + \frac{1}{4\pi \varepsilon_0} \frac{Q}{(r^2 + a^2 - 2ra \cos \theta)^{1/2}}.
\] (13.180)
CHAPTER 13. A PARTIAL DIFFERENTIAL EQUATION

Now we will apply the boundary conditions.

1. The electric potential must be finite as, \( r \to \infty \).

2. At any point on the surface of the sphere the electric potential must be zero since the sphere is grounded. That means

\[
V (R, \theta, \varphi) = 0.
\]  

(13.181)

Applying the first boundary condition \( V (r \to \infty, \theta, \varphi) \) must be finite (basically it must be zero since there are no other charges) gives

\[
\lim_{r \to \infty} V (r, \theta, \varphi) = \lim_{n \to \infty} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left( B_l r^l + \frac{C_l}{r^{l+1}} \right) P_l^m (\cos \theta) \\
\times (A_m \cos (m \varphi) + B_m \sin (m \varphi)) + \frac{1}{4 \pi \varepsilon_0} \frac{Q}{(r^2 + a^2 - 2ra \cos \theta)^{1/2}}.
\]  

(13.182)

since this expression diverges, we must have

\[
B_l = 0.
\]  

(13.183)

Therefore, we may rewrite the electric potential as

\[
V (r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{C_l}{r^{l+1}} P_l^m (\cos \theta) (A_m \cos (m \varphi) + B_m \sin (m \varphi)) \\
+ \frac{1}{4 \pi \varepsilon_0} \frac{Q}{(r^2 + a^2 - 2ra \cos \theta)^{1/2}}.
\]  

(13.184)

Now using the second boundary condition \( V (r = R, \theta, \varphi) = 0 \), we have

\[
V (R, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{C_l}{R^{l+1}} P_l^m (\cos \theta) (A_m \cos (m \varphi) + B_m \sin (m \varphi)) \\
+ \frac{1}{4 \pi \varepsilon_0} \frac{Q}{(R^2 + a^2 - 2Ra \cos \theta)^{1/2}} = 0.
\]  

(13.185)

Recall that Generating Function for the Legendry Polynomials

\[
\frac{1}{|r' - r|^2} = \frac{1}{\sqrt{r'^2 + r^2 - 2rr' \cos \theta}} = \sum_{l=0}^{\infty} \frac{r'_l}{r^{l+1}} P_l (\cos \theta)
\]  

(13.186)

we may write

\[
\frac{1}{4 \pi \varepsilon_0} \frac{Q}{(R^2 + a^2 - 2Ra \cos \theta)^{1/2}} = \frac{Q}{4 \pi \varepsilon_0} \sum_{l=0}^{\infty} \frac{R^l}{a^{l+1}} P_l (\cos \theta)
\]  

(13.187)
for $R < a$ and the electric potential on the surface of the sphere becomes

$$V(R, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{C_l}{R^{l+1}} P_l^m(\cos \theta) (A_m \cos (m\varphi) + B_m \sin (m\varphi)) + \frac{Q}{4\pi\epsilon_0} \sum_{l=0}^{\infty} \frac{R^l}{a^{l+1}} P_l(\cos \theta) = 0. \quad (13.188)$$

The above equation to be zero independent of $\varphi$, we at least have the coefficients for $P_l^m(\cos \theta)$ must be independent of $\cos (m\varphi)$ and $\sin (m\varphi)$. This requires

$$A_m = 0, B_m = 0 \text{ for } m \neq 0. \quad (13.189)$$

and the above expression becomes

$$V(R, \theta, \varphi) = \sum_{l=0}^{\infty} \frac{C_l}{R^{l+1}} P_l(\cos \theta) + \frac{Q}{4\pi\epsilon_0} \sum_{l=0}^{\infty} \frac{R^l}{a^{l+1}} P_l(\cos \theta) = 0. \quad (13.190)$$

$$V(R, \theta, \varphi) = 0 \Rightarrow \sum_{l=0}^{\infty} \left( \frac{C_l}{R^{l+1}} + \frac{Q}{4\pi\epsilon_0} \frac{R^l}{a^{l+1}} \right) P_l(\cos \theta) = 0. \quad (13.191)$$

There follows that

$$\frac{C_l}{R^{l+1}} + \frac{Q}{4\pi\epsilon_0} \frac{R^l}{a^{l+1}} = 0 \Rightarrow C_l = -\frac{Q}{4\pi\epsilon_0} \frac{R^{2l+1}}{a^{l+1}}. \quad (13.192)$$

Therefore, the electric potential is given by

$$V(r, \theta, \varphi) = -\frac{Q}{4\pi\epsilon_0} \sum_{l=0}^{\infty} \frac{1}{r^{l+1}} \frac{R^{2l+1}}{a^{l+1}} P_l(\cos \theta) + \frac{1}{4\pi\epsilon_0} \frac{Q}{(r^2 + a^2 - 2ra \cos \theta)^{1/2}}. \quad (13.193)$$

If we rewrite the series term as

$$-\frac{Q}{4\pi\epsilon_0} \sum_{l=0}^{\infty} \frac{1}{r^{l+1}} \frac{R^{2l+1}}{a^{l+1}} P_l(\cos \theta) = -\frac{Q}{4\pi\epsilon_0} \sum_{l=0}^{\infty} \frac{1}{r^{l+1}} \frac{R^{2l}}{a^l} P_l(\cos \theta)$$

$$= q' \frac{1}{4\pi\epsilon_0} \sum_{l=0}^{\infty} \frac{r^l}{r^{l+1}} P_l(\cos \theta) \quad (13.194)$$

with $r' = \frac{R^2}{a}$ and $q' = -\frac{Q R}{a}$, we can apply the generating function for Legendry polynomials

$$\frac{1}{|r' - r|^2} = \frac{1}{\sqrt{r'^2 + r^2 - 2rr' \cos \theta}} = \sum_{l=0}^{\infty} \frac{r^l}{r^{l+1}} P_l(\cos \theta) \quad (13.195)$$

and write

$$-\frac{Q}{4\pi\epsilon_0} \sum_{l=0}^{\infty} \frac{1}{r^{l+1}} \frac{R^{2l+1}}{a^{l+1}} P_l(\cos \theta) = + q' \frac{1}{4\pi\epsilon_0} \frac{1}{\sqrt{r^2 + r'^2 - 2rr' \cos \theta}}. \quad (13.196)$$
Then the electric potential can be expressed as

\[ V(r, \theta, \varphi) = \frac{q'}{4\pi \varepsilon_0} \frac{1}{\sqrt{r'^2 + r'^2 - 2rr' \cos \theta}} + \frac{Q}{4\pi \varepsilon_0} \frac{1}{\sqrt{r^2 + a^2 - 2ra \cos \theta}}, \]

(13.197)

where \( r' = \frac{R^2}{a} \) and \( q' = -Q \frac{R}{a} \). The above result shows that the electric potential due to a point charge \( Q \) placed outside the sphere at a distance \( a \) from the center of a grounded conducting sphere can be imagined as sum of the potential due to the point charge \( Q \) and a negative "image charge" \( q' = -Q \frac{R}{a} \) located inside the sphere at a distance \( r' = \frac{R^2}{a} \) from the center of the sphere. In reality this image charge is not a charge existing at this position. It is the total induced charge on the grounded conducting sphere due to the electric field produced by the point charge \( Q \).

The method of replacing this induced charge by an image charge to find the electric potential is called the method of images. It is a useful method in determining the electric potential of charges placed around conducting sphere, infinitely long conducting cylinder, or infinitely wide conducting plate.
Chapter 14

Functions of Complex Variables

14.1 Complex variables and functions

Real, Imaginary, and Complex Numbers: numbers that have the form

\[ z = a + ib, \quad (14.1) \]

with \( a \) and \( b \) real numbers and

\[ i = \sqrt{-1} \Rightarrow i^2 = -1 \quad (14.2) \]

are called complex numbers. In any complex number, \( z = a + ib \), \( a \) is the real part and \( b \) is the imaginary part

\[ \Re z = a, \Im z = b \quad (14.3) \]

Rectangular and Polar Representation of Complex Numbers: In the rectangular complex plane the real part \( a \) is the \( x \) coordinate and the imaginary part \( b \) is the \( y \) coordinate

\[ a = x, b = y \quad (14.4) \]

In polar coordinate the complex number \( z = a + ib \) is represented by

\[ a = r \cos \theta, b = r \sin \theta \quad (14.5) \]

where \( r \) is called the modulus or the magnitude and \( \theta \) is called the phase of the complex number \( z \)

Euler’s Equation and Exponential Representation of Complex Numbers: Any complex number \( z = x + iy \) can be expressed in an exponential form

\[ z = x + iy = r \exp(i \theta) = r \cos \theta + ir \sin \theta \quad (14.6) \]
CHAPTER 14. FUNCTIONS OF COMPLEX VARIABLES

\[ z = x + iy = r \cos \theta + ir \sin \theta \]

Figure 14.1: Rectangular and Polar representation of a complex number \( z = x + iy \).

where the magnitude and the phase are given by

\[ |z| = r = \sqrt{x^2 + y^2}, \theta = \tan^{-1} \left( \frac{y}{x} \right) \]  \hspace{1cm} (14.7)

Complex Conjugate and Magnitude of a Complex Number: The complex conjugate of a complex number \( z = x + iy \) is denoted by \( z^* \) and is given by

\[ z^* = x - iy \]  \hspace{1cm} (14.8)

The magnitude of a complex number \( z \) is related to its complex conjugate by

\[ |z| = \sqrt{zz^*} = \sqrt{x^2 + y^2} \]  \hspace{1cm} (14.9)

Example 13.1 Consider the complex function

\[ f(z) = \frac{2z + 1}{z - i} \]  \hspace{1cm} (14.10)

Find the real and imaginary parts of this function.

Solution: To find the real and imaginary part of a complex function for the complex variable

\[ z = x + iy \]  \hspace{1cm} (14.11)

we must be able to express the function \( f(z) \) as

\[ f(x + iy) = u(x, y) + iv(x, y). \]  \hspace{1cm} (14.12)
In order to do that we substitute, $z = x + iy$, into the given function

$$f(x + iy) = \frac{2(x + iy) + 1}{(x + iy) - i} = \frac{2x + 1 + i2y}{x + i(y - 1)} = \frac{[2x + 1 + i2y][x - i(y - 1)]}{x^2 + (y - 1)^2}$$

$$= \frac{(2x + 1)x + 2y(y - 1) + i[2xy -(2x + 1)(y - 1)]}{x^2 + (y - 1)^2}$$

$$\Rightarrow f(x + iy) = \frac{(2x + 1)x + 2y(y - 1) + i[-y + 2x + 1]}{x^2 + (y - 1)^2}$$ (14.13)

There follows that

$$u(x, y) = \frac{(2x + 1)x + 2y(y - 1)}{x^2 + (y - 1)^2}, v(x, y) = \frac{2x - y + 1}{x^2 + (y - 1)^2}$$ (14.14)

Example 13.2 Find the real and imaginary parts of the complex function

$$f(z) = z^{1/4}.$$ (14.15)

Solution: Whenever we are given exponential complex functions it is better to use polar coordinates. Let

$$z = r \exp(i\theta),$$ (14.16)

where $r$ and $\theta$ are related to Cartesian coordinates $x$ and $y$ by

$$r = \sqrt{x^2 + y^2}, \theta = \tan^{-1}\left(\frac{y}{x}\right).$$ (14.17)

The function can then be expressed as

$$f(z) = [r \exp(i\theta)]^{1/4} = r^{1/4} \exp\left(\frac{i\theta}{4}\right) = r^{1/4} \cos\left(\frac{\theta}{4}\right) + ir^{1/4} \sin\left(\frac{\theta}{4}\right)$$ (14.18)

so that the real and imaginary parts can be expressed as

$$u(r, \theta) = r^{1/4} \cos\left(\frac{\theta}{4}\right), \quad v(r, \theta) = r^{1/4} \sin\left(\frac{\theta}{4}\right)$$ (14.19)

14.2 Analytic Functions

A function $f(x)$ is analytic (or regular or holomorphic or monogenic) in a region of the complex plane if it has a (unique) derivative at every point of the region. The statement "$f(z)$ is analytic at a point $z = z_0$" means that $f(z)$ has a derivative at every point inside some small circle about $z = z_0$.

$$\frac{df}{dz}|_{z = z_0} = \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}$$ (14.20)
While the functions
\[
\frac{d}{dz} (z^2) = 2z, \quad \frac{d}{dz} (\cos z) = -\sin z, \quad \frac{d}{dz} \left( \sqrt{1 + z^2} \right) = \frac{z}{\sqrt{1 + z^2}}
\]
are analytic, the function
\[
f(z) = |z|^2 = \sqrt{x^2 + y^2}
\]
is not. So how one can determine whether a given function of complex variable is analytic or not. We use *The Cauchy-Riemann Conditions*. It states that if the complex function
\[
f(z) = u(x, y) + iv(x, y),
\]
where \(u(x, y)\) and \(v(x, y)\) are real functions, is analytic in a region, \(R\), then in that region the following two conditions must be satisfied
\[
\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.
\]
Next we shall derive these conditions. To this end, we note that for \(z = x + iy\), we can write
\[
\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}, \quad \frac{\partial f}{\partial y} = \frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y}
\]
and substituting these expressions into
\[
\frac{\partial f}{\partial x} = -i \frac{\partial f}{\partial y}
\]
on the other hand using \(f(z) = u(x, y) + iv(x, y)\), we may also write
\[
\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}, \quad \frac{\partial f}{\partial y} = \frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y}
\]
and substituting these expressions into
\[
\frac{\partial f}{\partial x} = -i \frac{\partial f}{\partial y}
\]
one finds
\[
\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = -i \left( \frac{\partial u}{\partial y} + i \frac{\partial v}{\partial y} \right) \Rightarrow \frac{\partial u}{\partial x} + \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}.
\]
Two complex function are equal if and only if their real part are equal and their imaginary part are equal. Therefore,
\[
\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y},
\]
which is *The Cauchy-Riemann Conditions*. 
14.3 Important Terminologies

The following terminologies are important in the description of functions of complex variable.

(a) **Regular point:** The point $z_o$ is said to be a regular point of the function $f(z)$ if the function $f(z)$ is analytic at that point.

(b) **Singular point:** If $f(z)$ is not analytic at a point, $z_o$, then $z_o$ is said to be a singular point, or singularity of $f(z)$.

(c) **Isolated Singularity:** If $z_o$ is the only singular point for the function $f(z)$ within an arbitrarily small region surrounding $z_o$, then $z_o$ is said to be an isolated singular point for the function $f(z)$.

(d) **Harmonic and conjugate Harmonic Functions:** Consider the first Cauchy-Riemann Condition

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}. \quad (14.31)$$

Upon differentiating with respect to $x$, we find

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y}. \quad (14.32)$$

Similarly, differentiate the second condition

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \quad (14.33)$$

with respect to $y$, we have

$$\frac{\partial^2 v}{\partial x \partial y} = -\frac{\partial^2 u}{\partial y^2} \Rightarrow \frac{\partial^2 u}{\partial y^2} = -\frac{\partial^2 v}{\partial x \partial y}. \quad (14.34)$$

so that adding the two equations, one finds

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 v}{\partial x \partial y} - \frac{\partial^2 v}{\partial x \partial y} = 0 \Rightarrow \nabla^2 u = 0 \quad (14.35)$$

which is Laplace’s equation in two dimension. Functions satisfying Laplace’s equation are called harmonic functions. Thus solutions of Laplace’s equation which are real and imaginary parts of a function $f(z)$ are called conjugate harmonic functions.

An Important Theorem: Let $f(z)$ be analytic within a region $R$, with one or more singular points on the borders (and possibly outside) of a region, $R$. Then $f(z)$ has derivatives of all orders that exist at any point $z_k$ inside $R$, so that $f(z)$ can be expanded in a Taylor series about the point $z_k$. This series converges for all points within a circle centered at $z_1$ that extends to the closest singularity of $f(z)$. 


Example 13.3 Expand the complex function given below in a Taylor series about the origin, and find the circle of convergence for this series.

\[ f(z) = \sqrt{z + 2i} \quad (0 \leq \theta \leq 2\pi) \]  

(14.36)

Solution: The Taylor series for a function, \( f(z) \), is given by

\[ f(z) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{d^n f(z)}{dz^n} \bigg|_{z=z_0} (z-z_0)^n. \]  

(14.37)

For the complex function, \( f(z) = \sqrt{z + 2i} \), we have

\[
\begin{align*}
\frac{df(z)}{dz} &= \frac{1}{2\sqrt{z + 2i}} \Rightarrow f(z_0) = \sqrt{z_0 + 2i} \\
\frac{d^2 f(z)}{dz^2} &= \frac{1}{2^2 (z + 2i)^{3/2}} \Rightarrow \frac{d^2 f(z)}{dz^2} \bigg|_{z=z_0} = \frac{1}{2^2 (z_0 + 2i)^{3/2}} \\
\frac{d^3 f(z)}{dz^3} &= \frac{3}{2^3 (z + 2i)^{5/2}} \Rightarrow \frac{d^3 f(z)}{dz^3} \bigg|_{z=z_0} = \frac{1}{2^3 (z_0 + 2i)^{5/2}}
\end{align*}
\]

(14.38)

Therefore

\[
\begin{align*}
f(z) &= \sqrt{z_0 + 2i} + \frac{1}{2} \frac{1}{(z_0 + 2i)^{1/2}} (z-z_0) - \frac{1}{2^3} \frac{1}{(z_0 + 2i)^{3/2}} (z-z_0)^2 \\
&\quad + \frac{1}{2^4} \frac{1}{(z_0 + 2i)^{5/2}} (z-z_0)^3 \ldots
\end{align*}
\]

(14.39)

This series diverges at \( z_0 = -2i \),

\[
\lim_{z_0 \to -2i} \frac{1}{z_0 + 2i} \to \infty
\]

which shows that the function \( f(z) \) has a singular point at \( z_0 = -2i \). Therefore, the radius of convergence is

\[ |z_0| = 2. \]

Which means the series is convergent for all \( z \) inside the circle of radius, \( R = 2 \) centered about the origin, \( z_1 = 0 \).

Example 13.4 Polarization of EM waves (Extra Example): Consider an EM wave propagating along the positive z-direction. We can express the electric field of this wave using complex functions as

\[
\vec{E} = (E_{0x}\hat{x} + E_{0y}\hat{y}) e^{i(kz-\omega t)},
\]

(14.40)

where \( k \) is the wave number, \( \omega \) is the angular frequency, \( E_{0x} \) and \( E_{0y} \) are the amplitude of the electric field in the \( x \) and \( y \) direction, respectively.
14.3. IMPORTANT TERMINOLOGIES

(a) Express the complex electric field using Euler’s formula assuming $E_{0x}$ and $E_{0y}$ are complex.

(b) Assume the phase of the $x$ component $\varphi_x = 0$ and that of the $y$ component $\varphi_y = \varphi$, find the components of the electric field for $\varphi = 0, \pm \pi/2$.

(c) Find the real part of the Electric field vector for $\varphi = 0, \pm \pi/2$ and plot the corresponding results for $E_y$ vs $E_x$ for a fixed value of $z$ (for example $z = 0$) and $0 \leq \omega t \leq 2\pi$.

Solution:

(a) For complex electric field, using Euler’s formula, we may write

$$E_{0x} = |E_{0x}| e^{i\varphi_x}, E_{0y} = |E_{0y}| e^{i\varphi_y}$$

so that

$$\vec{E} = (E_{0x}\hat{x} + E_{0y}\hat{y}) e^{i(kz-\omega t)},$$

becomes

$$\vec{E} = (|E_{0x}| e^{i\varphi_x}\hat{x} + |E_{0y}| e^{i\varphi_y}\hat{y}) e^{i(kz-\omega t)} = |E_{0x}| e^{i(kz-\omega t+\varphi_x)\hat{x}} + |E_{0y}| e^{i(kz-\omega t+\varphi_y)\hat{y}}.$$  \hfill (14.43)

(b) If we assume the phase of the $x$ component $\varphi_x = 0$ and that of the $y$ component $\varphi_y = \varphi$, as measured relative to the phase of the $x$ component of the electric field, then we can write

$$\vec{E} = |E_{0x}| e^{i(kz-\omega t)\hat{x}} + |E_{0y}| e^{i(kz-\omega t+\varphi)\hat{y}}$$ \hfill (14.44)

If the phase $\varphi = 0$, the electric field becomes

$$\vec{E} = |E_{0x}| e^{i(kz-\omega t)\hat{x}} + |E_{0y}| e^{i(kz-\omega t)\hat{y}}$$ \hfill (14.45)

which leads to

$$E_y = |E_{0x}| e^{i(kz-\omega t)} = |E_{0x}| \cos (kz - \omega t) + i |E_{0x}| \sin (kz - \omega t),$$ \hfill (14.46)

$$E_x = |E_{0y}| e^{i(kz-\omega t)} = |E_{0y}| \cos (kz - \omega t) + i |E_{0y}| \sin (kz - \omega t).$$ \hfill (14.47)

For $\varphi = \pm \pi/2$, we still have the same form for the $x$-component

$$E_x = |E_{0y}| e^{i(kz-\omega t)} = |E_{0y}| \cos (kz - \omega t) + i |E_{0y}| \sin (kz - \omega t).$$ \hfill (14.48)

but for the $y$-component we find

$$E_y = |E_{0x}| e^{i(kz-\omega t)} = |E_{0x}| \cos (kz - \omega t) + i |E_{0x}| \sin (kz - \omega t),$$

$$E_y = |E_{0y}| e^{i(kz-\omega t)} = |E_{0y}| \cos (kz - \omega t \pm \pi/2) + i |E_{0y}| \sin (kz - \omega t \pm \pi/2).$$ \hfill (14.49)
Using
\[
\cos (\alpha \pm \pi/2) = \cos (\alpha) \cos (\pi/2) \mp \sin (\alpha) \sin (\pi/2) = \mp \sin (\alpha) \quad (14.50)
\]
and
\[
\sin (\alpha \pm \pi/2) = \sin (\alpha) \cos (\pi/2) \pm \cos (\alpha) \sin (\pi/2) = \pm \cos (\alpha) \quad (14.51)
\]
we may write the y-component of the electric field as
\[
E_y = \mp |E_{0y}| \sin (kz - \omega t) \pm i |E_{0y}| \cos (kz - \omega t). \quad (14.52)
\]
and light is linearly polarized. Under this condition if \( |E_{0y}| = 0 \) the light is x-polarized; and if \( |E_{0x}| = 0 \), the light is y-polarized.

\( (c) \) For \( \varphi = 0 \ (z = 0) \) using the results above, we may write
\[
E_x = |E_{0x}| \cos (-\omega t) + i |E_{0x}| \sin (-\omega t) = |E_{0x}| \cos (\omega t) - i |E_{0x}| \sin (\omega t), \quad (14.53)
\]
\[
E_y = |E_{0y}| \cos (-\omega t) + i |E_{0y}| \sin (-\omega t) = |E_{0y}| \cos (\omega t) - i |E_{0y}| \sin (\omega t).
\]
and the real electric field vector would be
\[
\text{Re} \left( \vec{E} \right) = \text{Re} (E_x) \hat{x} + \text{Re} (E_y) \hat{y} = |E_{0x}| \cos (\omega t) \hat{x} + |E_{0y}| \cos (\omega t) \hat{y} \quad (14.54)
\]
Plotting the \( y \) vs the \( x \) components for different we find
This kind of EM wave (light) is called \textit{Linearly (plane) polarized} EM (light).

For $\varphi = \pm \pi/2$ ($z = 0$) using the results in part b,

\begin{equation}
E_x = |E_{0x}| \cos(\omega t) - i |E_{0x}| \sin(\omega t),
\end{equation}

\begin{equation}
E_y = \mp |E_{0y}| \sin(-\omega t) \pm i |E_{0y}| \cos(-\omega t) = \pm |E_{0y}| \sin(\omega t) \pm i |E_{0y}| \cos(\omega t)
\end{equation}

\Rightarrow

\begin{equation}
E_y = \pm (|E_{0y}| \sin(\omega t) + i |E_{0y}| \cos(\omega t)).
\end{equation}
and the real electric field vector would be

$$\text{Re} \left( \vec{E} \right) = \text{Re} \left( E_x \right) \hat{x} + \text{Re} \left( E_y \right) \hat{y} = |E_{0x}| \cos(\omega t) \hat{x} \pm |E_{0y}| \sin(\omega t) \hat{y}$$

(14.57)

Plotting the $y$ vs the $x$ components for different $\omega$ we find

```
In[18]:  E0x = 1
E0y = 2
w = 2 \pi

Out[18]= 1
Out[17]= 2
Out[16]= 2 \pi
```

```
In[10]:  ParametricPlot[{E0x Cos[w t], E0y Sin[w t]}, {t, 0, 1}]
```

This kind of EM wave (light) is called Elliptically (left or right) polarized EM (light). If $|E_{0x}| = |E_{0y}|$, is is called Circularly (left or right) polarized.

### 14.4 Contour Integration and Cauchy’s Theorem

**Cauchy’s theorem:** Consider a function, $f(z)$, that is analytic inside a region, $R$, on the complex plane. Within the region, $R$, if there is a closed curve $C$ that does not cross itself, and that has a finite number of sharp corners, then
Cauchy’s theorem states that
\[ \oint_C f(z) \, dz = 0. \] (14.58)
The line integral on this curve is always zero.

**Proof:** Using
\[ z = x + iy \] (14.59)
for a complex function
\[ f(z) = u(x, y) + iv(x, y) \] (14.60)
the closed integral can be expressed as
\[
\oint_C f(z) \, dz = \oint_C \left[ (u(x, y) + iv(x, y)) \, (dx + idy) \right]
\]
\[
= \oint_C \left[ (u(x, y) dx - v(x, y) dy) + i \oint_C \left[ (v(x, y) dx + u(x, y) dy) \right] \right] \] (14.61)
Applying Green’s theorem [that you were introduced to last semester]
\[
\oint_C [(P(x, y) \, dx + Q(x, y) \, dy)] = \iint_A \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \, dx \, dy. \] (14.62)
we may write
\[ \oint_C f(z)\,dz = \oint_C [(u(x, y)\,dx - v(x, y)\,dy)] + i \oint_C [(v(x, y)\,dx + u(x, y)\,dy)] \]
\[ = \iint_A \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \,dxdy + i \iint_A \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \,dxdy \]  
(14.63)
\[ = -\iint_A \left( \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \,dxdy + i \iint_A \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) \,dxdy. \]  
(14.64)

For analytic function,
\[ f(z) = u(x, y) + iv(x, y), \]  
(14.65)
the Cauchy-Riemann condition states that
\[ \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \]  
(14.66)

Applying these conditions, one can easily finds
\[ \oint_C f(z)\,dz = -\iint_A \left( \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \,dxdy + i \iint_A \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) \,dxdy = 0. \]  
(14.67)

The Equation of a Circle on the Complex Plane: Consider a circle on a complex plane centered about a point, \( z_0 = x_0 + iy_0 \), shown in Fig. 14.2.

Any point on this circle, \( z = x + iy \), can be described by the equation
\[ z - z_0 = (x - x_0) + i(y - y_0) = |z - z_0|e^{i\theta}. \]  
(14.68)
where
\[ |z - z_0| = \sqrt{(x - x_0)^2 + (y - y_0)^2}, \theta = \tan^{-1} \left( \frac{y - y_0}{x - x_0} \right). \]

14.5 Cauchy’s Integral formula

Suppose the complex function, \( f(z) \), is analytic inside a region enclosed by a curve \( C \) and also on the curve itself (e.g. Fig. 14.3). The value for this function at a point, \( z = z_0 \), inside the curve \( C \), \( f(z_0) \) can be determined using the line integral on the closed curve, \( C \), given by
\[ f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0} \,dz. \]  
(14.69)
14.5. CAUCHY’S INTEGRAL FORMULA

Figure 14.2: A circle centered about $z_0$ on a complex plane.

Proof: Consider the function

$$g(z) = \frac{f(z)}{z - z_0}$$

which is analytic everywhere inside the closed curve, $C$, except at $z = z_0$. To find the integral of this function over the given closed curve, $C$, we first exclude this singular point from the region bounded by the curve $C$. In Fig. 14.4 we have formed a region enclosed by a new closed curve, $C'$, that consist of four different curves ($C_1, C_2, l_1, l_2$) that excluded the singular point, $z = z_0$. Everywhere in this region, the function $g(z)$ is analytic as the only singular point, $z = z_0$ is no longer in the region. Then according to Cauchy’s Theorem, we must have

$$I = \oint_{C'} g(z) \, dz = \frac{1}{2\pi i} \int_{C'} \frac{f(z)}{z - z_0} \, dz = 0,$$  \hspace{1cm} (14.71)

where $C'$ is the curve shown in the Fig. 14.4. This integral can be split into four integrals along four different curves ($C_1, C_2, l_1, l_2$),

$$\oint_{C'} g(z) \, dz = \int_{C_1} g(z) \, dz + \int_{l_2} g(z) \, dz + \int_{C_2} g(z) \, dz + \int_{l_1} g(z) \, dz = 0.$$  \hspace{1cm} (14.72)
Since we are interested in the integral
\[ I = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0} \, dz = \lim_{\epsilon \to 0} \int_{C_1} g(z) \, dz, \] (14.73)
in the limit \( \epsilon \to 0 \), integrating along \( l_1 \) and \( l_2 \) leads to same function integrated in opposite direction over the same line, one finds
\[ \int_{l_1} g(z) \, dz = -\int_{l_2} g(z) \, dz. \] (14.74)

Therefore, Eq. (??) becomes
\[ \lim_{\epsilon \to 0} \oint_{C_1} g(z) \, dz = \lim_{\epsilon \to 0} \left[ \oint_{C_1} g(z) \, dz + \oint_{C_2} g(z) \, dz \right] = 0 \] (14.75)
which gives
\[ \lim_{\epsilon \to 0} \oint_{C_1} \frac{f(z)}{z - a} \, dz = -\lim_{\epsilon \to 0} \oint_{C_2} \frac{f(z)}{z - a} \, dz. \] (14.76)
Note that from Fig. 14.4 the integration over $C_1$ is counterclockwise where as the integration over $C_2$ is clockwise. Now using equation of a circle on a complex plane we saw earlier, we may write

$$z - z_0 = e e^{i\theta} \Rightarrow z = z_0 + e e^{i\theta} \Rightarrow dz = i e e^{i\theta} d\theta$$  \hspace{1cm} (14.77)

so that

$$\lim_{\epsilon \to 0} \oint_{C_1} \frac{f(z)}{z-a} \, dz = - \lim_{\epsilon \to 0} \oint_{C_2} \frac{f(z)}{z-a} \, dz = - \lim_{\epsilon \to 0} \oint_{C_1} \frac{f(z_0 + e e^{i\theta})}{e e^{i\theta}} i e e^{i\theta} d\theta$$

$$\Rightarrow \lim_{\epsilon \to 0} \oint_{C_1} \frac{f(z)}{z-a} \, dz = \lim_{\epsilon \to 0} \int_{0}^{2\pi} f(z_0 + e e^{i\theta}) \, d\theta = f(z_0) \int_{0}^{2\pi} d\theta$$

$$\Rightarrow \lim_{\epsilon \to 0} \oint_{C_1} \frac{f(z)}{z-a} \, dz = 2\pi i f(z_0)$$  \hspace{1cm} (14.78)

Thus in this limit $\epsilon \to 0$, one can write

$$\oint_{C} \frac{f(z)}{z-z_0} \, dz = \lim_{\epsilon \to 0} \oint_{C_1} \frac{f(z)}{z-a} \, dz = i \int_{0}^{2\pi} f(z_0) \, d\theta = 2\pi i f(z_0)$$

$$\Rightarrow f(z_0) = \frac{1}{2\pi i} \oint_{C} \frac{f(z)}{z-z_0} \, dz.$$  \hspace{1cm} (14.79)
Note: A line integral in the complex plane is called a contour integral.

Example 13.5 Evaluate the contour integral

$$I = \int_{1+i}^{2+4i} z^2 dz$$ (14.80)

along the paths indicated.

(a) A straight line joining the points $z_1 = 1 + i$ and $z_2 = 2 + 4i$.

(b) Two straight lines: the first from the point $z_1 = 1 + i$ to $z_0 = 2 + i$, and the second from the point $z_0$ to the point $z_2 = 2 + 4i$.

(c) Find the integral

$$I = \oint_C z^2 dz$$ (14.81)

for the closed triangular curve

![Figure 14.5: Contour integration.](image)

Solution:
(a) First, we just simply carry out the complex function integral

\[ \int_{z_1}^{z_2} z^2 \, dz = \frac{z^3}{3} \bigg|_{z_1}^{z_2} = \frac{z_2^3 - z_1^3}{3} \quad = \frac{(2 + 4i)^3 - (1 + i)^3}{3} \quad = -\frac{86}{3} - 6i \quad (14.82) \]

We now integrate the line joining the two points \( z_1 = 1 + i \) and \( z_2 = 2 + 4i \) on the complex plane. The equation for the line joining the two points

\[ \frac{y - y_1}{x - x_1} = \text{slope} \Rightarrow \frac{y - 1}{x - 1} = 3 \Rightarrow y = 3x - 2. \quad (14.83) \]

Using

\[ z = x + iy \quad (14.84) \]

one can write the integral

\[ \int z^2 \, dz = \int \left( x^2 - y^2 + 2ixy \right) (dx + idy) \]

\[ = \int \left[ (x^2 - y^2) \, dx + i \left( (x^2 - y^2) + 2xy \right) \, dy \right] \]

\[ \Rightarrow \int z^2 \, dz = \int f(x, y) \, dx + i \int f(x, y) \, dy, \quad (14.85) \]

where

\[ f(x, y) = x^2 - y^2 + 2ixy \quad (14.86) \]

Using the equation of the line joining the two points, we have

\[ y = 3x - 2 \Rightarrow dy = 3dx, \]

\[ f(x, y) = x^2 - (3x - 2)^2 + 2ix(3x - 2) \]

\[ = -8x^2 + 12x - 4 + 2i(3x^2 - 2x) \quad (14.87) \]

so that

\[ \int_1^2 f(x, y) \, dx = \int_1^2 \left[ -8x^2 + 12x - 4 + 2i(3x^2 - 2x) \right] \, dx \]

\[ \Rightarrow \int_1^2 f(x, y) \, dx = -\frac{8}{3}x^3 + 6x^2 - 4x + 2i \left( x^3 - x^2 \right) \bigg|_1 \]

\[ \Rightarrow \int_1^2 f(x, y) \, dx = -\frac{14}{3} + 8i \]

\[ \int_1^2 f(x, y) \, dy = 3 \int_1^2 \left[ -8x^2 + 12x + 4 + 2i(3x^2 - 2x) \right] \, dx \]

\[ \Rightarrow \int_1^2 f(x, y) \, dy = 3 \left( -\frac{14}{3} + 8i \right) = -14 + 24i \]
and
\[
\int_{z_1}^{z_2} z^2 dz = \int_1^2 f(x, y) \, dx + i \int_1^2 f(x, y) \, dy = -\frac{14}{3} + 8i + i(-14 + 24i)
\]
\[
\Rightarrow \int_{z_1}^{z_2} z^2 dz = -\frac{86}{3} - 6i
\]

(b) The integral along the first line from the point \(z_1 = 1 + i\) to \(z_0 = 2 + i\), and the second line from the point \(z_0\) to the point \(z_2 = 2 + 4i\) can be expressed as
\[
\int_{z_1}^{z_2} z^2 dz = \int_{z_0}^{z_2} z^2 dz = \int_{z_0}^{z_1} z^2 dz + \int_{z_2}^{z_0} z^2 dz = \int_{z_1}^{z_0} z^2 dz + \int_{z_2}^{z_1} z^2 dz
\]
\[
= -\frac{86}{3} - 6i
\]

or following the two path
\[
\int_{1+i}^{2+4i} z^2 dz = \int_1^2 (x^2 - 1 + 2ix) \, dx + \int_1^4 (2^2 - y^2 + 4iy) \, idy
\]
\[
= \left[ \frac{x^3}{3} + i \frac{x^2}{2} - x \right]_1^2 + i \left[ -\frac{y^3}{3} + 2iy^2 + 4y \right]_1^4
\]
\[
= \frac{4}{3} + 3i - 30 - 9i
\]
\[
\Rightarrow \int_{1+i}^{2+4i} z^2 dz = -\frac{86}{3} - 6i
\]

which is the same result as in part a.

(c) Using Cauchy’s theorem
\[
\oint_C f(z) \, dz = 0
\]

for the function
\[
f(z) = z^2
\]

which is analytic everywhere inside and on the triangle, we can easily find
\[
\oint_C z^2 dz = 0.
\]

Noting that
\[
\oint_C z^2 dz = \int_{z_1}^{z_0} z^2 dz + \int_{z_0}^{z_2} z^2 dz + \int_{z_2}^{z_1} z^2 dz = \int_{z_1}^{z_0} z^2 dz + \int_{z_0}^{z_2} z^2 dz - \int_{z_1}^{z_0} z^2 dz
\]

and using the result we obtained in part a and b,
\[
\oint_C z^2 dz = -\frac{86}{3} - 6i - \left( -\frac{86}{3} - 6i \right) = 0.
\]
Example 13.6 Evaluate the contour integral
\[ I = \oint \frac{e^z}{z(z+1)} \, dz \] (14.96)
for the closed path \( C \) given by \(|z-1|=3\).

\[ \text{Solution:} \] Noting that
\[ \frac{1}{z(z+1)} = \frac{1}{z} - \frac{1}{z+1} \] (14.97)
we may write
\[ I = \oint \frac{e^z}{z} \, dz = \oint \frac{e^z}{z+1} \, dz + \oint \frac{e^z}{z+1} \, dz \] (14.98)
For the contour defined by the curve \(|z-1|=3\) is shown in Fig. 14.6. Noting that the function
\[ f(z) = e^z \] (14.99)
is analytic inside and on the curve bounded by \( C \) and \( z=0 \) and \( z=-1 \) are inside this curve, applying Cauchy’s integral formula we can write
\[ \oint \frac{e^z}{z} \, dz = 2\pi i f (0) = 2\pi i, \quad \oint \frac{e^z}{z+1} \, dz = 2\pi i f (-1) = 2\pi i e^{-1} \] (14.100)
so that
\[ I = \oint \frac{e^z}{z(z+1)} \, dz = 2\pi i (1 + e^{-1}) . \] (14.101)
14.6 Laurent’s Theorem

Consider a region $R$ between two circles $C_1$ and $C_2$ centered at the same point, $z = z_0$, and let $f(z)$ be analytic in $R$.

Then $f(z)$ can be expanded in a series given by

$$f(z) = a_0 + a_1 (z - z_0) + a_2 (z - z_0)^2 + \ldots + \frac{b_1}{z - z_0} + \frac{b_2}{(z - z_0)^2} + \ldots$$

$$= \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n},$$

that converges for any value of $z$ within the region $R$.

Some important terminologies

(a) A regular point: If $b_n = 0$ for all values of $n$, then $f(z)$ is analytic at $z = z_0$, and $z_0$ is said to be a regular point of $f(z)$.

(b) The residue: the coefficient $b_1$ in Laurent’s series is called the residue of $f(z)$.

(c) A pole: If the principal part of the series has terms only up to $b_N$ (i.e. $b_n = 0$ for all $n > N$), then $f(z)$ is said to have a pole of order $N$ at the point $z = z_0$

$$\sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n} = \frac{b_1}{z - z_0} + \frac{b_2}{(z - z_0)^2} + \ldots + \frac{b_N}{(z - z_0)^N}$$

where

$$b_{N+1} = b_{N+2} = b_{N+3} = \ldots = 0$$

(14.103)
(d) A simple pole: If the principal part of the series has only the single term $b_1$ (i.e. $b_{n+1} = 0$ for all $n \geq 1$), then $f(z)$ is said to have a simple pole at $z = z_0$.

(e) Essential singularity: If there are an infinite number of $b_n$'s different from zero, $f(z)$ has an essential singularity at $z = z_0$.

14.7 Residues and the Residue Theorem

The Residue Theorem: Let $z_k$ be an isolated singular point of $f(z)$ inside a closed curve defined by $C$, the residue theorem states that

$$\int_C f(z) \, dz = 2\pi i R(z_k),$$

where the integral is in a counterclockwise direction and $R(z_k)$ is the residue of the function $f(z)$ at $z = z_k$.

Proof: If we define a small circle of radius $r = \epsilon$, as shown in the figure below we can write [applying the procedure we followed when we proof the Cauchy integral formula, see Fig. 14.4] that

$$\int_C f(z) \, dz = -\int_{C_2} f(z) \, dz. \quad (14.104)$$
We can expand the function \( f(z) \) in Laurent series in the shaded region since the function is analytic in this region
\[
\begin{align*}
f(z) &= a_0 + a_1 (z - z_k) + a_2 (z - z_k)^2 + \ldots + \frac{b_1}{z - z_k} \\
&\quad + \frac{b_2}{(z - z_k)^2} + \ldots = \sum_{n=0}^{\infty} a_n (z - z_k)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_k)^n}
\end{align*}
\]
(14.105)

This leads to
\[
\begin{align*}
\oint_C f(z) \, dz &= a_0 \oint_{C_2} dz + a_1 \oint_{C_2} (z - z_k) \, dz + a_2 \oint_{C_2} (z - z_k)^2 \, dz \\
&\quad + \ldots + \frac{b_1}{\oint_{C_2} (z - z_k)} \, dz + \frac{b_2}{\oint_{C_2} (z - z_k)^2} \, dz + \frac{b_3}{\oint_{C_2} (z - z_k)^3} \, dz \ldots 
\end{align*}
\]
(14.106)

Since the function \( (z - z_0)^n \) is analytic everywhere in the region bounded by the curve \( C_2 \) for \( n \geq 0 \), we have
\[
\oint_{C_2} (z - z_k)^n \, dz = 0,
\]
(14.107)

so that
\[
\oint_C f(z) \, dz = b_1 \oint_{C_2} \frac{1}{z - z_k} \, dz + b_2 \oint_{C_2} \frac{1}{(z - z_k)^2} \, dz + b_3 \oint_{C_2} \frac{1}{(z - z_k)^3} \, dz \ldots 
\]
(14.108)

Now using
\[
z - z_k = e^{i\theta} \Rightarrow z = z_k + e^{i\theta} \Rightarrow dz = ie^{i\theta} d\theta
\]
(14.109)

we have
\[
\oint_C f(z) \, dz = b_1 \oint_{C_2} \frac{1}{e^{i\theta}} ie^{i\theta} d\theta + b_2 \oint_{C_2} \frac{1}{(e^{i\theta})^2} ie^{i\theta} d\theta + b_3 \oint_{C_2} \frac{1}{(e^{i\theta})^3} ie^{i\theta} d\theta \ldots 
\]
(14.110)

which may want to put in the form
\[
\oint_C f(z) \, dz = 2\pi i b_1 + \sum_{n=2}^{\infty} \frac{ib_n}{\int_{C_2} e^{-i(n-1)\theta}} d\theta
\]
\[
= 2\pi i b_1 + \sum_{n=2}^{\infty} \frac{ib_n}{\int_{C_2}} e^{-i(n-1)\theta} d\theta
\]
\[
= 2\pi i b_1 + \sum_{n=2}^{\infty} \frac{ib_n}{\int_{C_2}^{2\pi}} \{ \cos [(n - 1) \theta] - i \sin [(n - 1) \theta] \} d\theta
\]
(14.111)

Since
\[
\int_0^{2\pi} \cos [(n - 1) \theta] \, d\theta = -i \int_0^{2\pi} \sin [(n - 1) \theta] \, d\theta = 0
\]
(14.112)

for all \( n \geq 2 \), we find
\[
\oint_C f(z) \, dz = 2\pi i b_1 = 2\pi i R(z_k),
\]
(14.113)
where
\[ R(z_k) = b_1 \tag{14.114} \]

is the residue of the function \( f(z) \) at \( z = z_k \).

**Note:** If there are more than one residue, the residue theorem becomes
\[
\oint f(z) \, dz = 2\pi i \sum_{k=1}^{n} R(z_k). \tag{14.115}
\]

### 14.8 Methods of finding residues

**Method 1:** Expand \( f(z) \) about the point \( z_0 \) and simply read the value of the residue off of the series. Recall Taylor Series expansions for some basic functions (PHYS 3150): Any function, \( f(x) \) that is differentiable for all values of \( x \) in the specified domain, can be expressed in Taylor or Maclaurin series. That means
\[
\frac{d^n f(x)}{dx^n} \text{ exists for } n \geq 0 \text{ and } x \in \text{Re domain}
\]

\( \Rightarrow f(x) = \sum_{n=0}^{\infty} b_n (x - a)^n, \text{ where } b_n = \frac{1}{n!} \frac{d^n f(x)}{dx^n} \bigg|_{x=a} \tag{14.116} \)

This can be applied for complex \( z \). Taylor Series for some common functions:

\[
sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n-1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \ldots \text{ convergent for all } x
\]

\[
\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \ldots \text{ convergent for all } x
\]

\[
e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \ldots \text{ convergent for all } x
\]

\[
\ln(1 + x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \ldots \text{ convergent for } -1 < x \leq 1;
\]

\[
(1 + x)^p = \sum_{n=1}^{\infty} \binom{p}{n} x^n = 1 + px + \frac{p(p-1)}{2!} x^2 + \frac{p(p-1)(p-2)}{3!} x^3 \ldots
\]

convergent for all \( |x| < 1 \) \tag{14.117}

where
\[
\binom{p}{n} = \frac{p(p-1)(p-2)(p-3)\ldots(p-n+1)}{n!}, \tag{14.118}
\]
is called binomial coefficient.

Power series expansions of functions are unique. That is, if you use different methods for finding a power-series expansion of a function, then you must get the same result. A power series expansion for a function must be the power series expansion for that function!
Example 13.7 Find $R(0)$ for the complex function

$$f(z) = z \cos \left(\frac{1}{z}\right)$$  \hfill (14.119)

**Solution:** We recall that the series expansion of the function

$$f(z) = \cos(z)$$  \hfill (14.120)

is

$$\cos(z) = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} \ldots$$  \hfill (14.121)

so that

$$\cos\left(\frac{1}{z}\right) = 1 - \frac{1}{z^2} + \frac{1}{z^4} - \frac{1}{z^6} \ldots$$  \hfill (14.122)

and

$$f(z) = z \cos\left(\frac{1}{z}\right) = z - \frac{1}{z} + \frac{1}{z^3} - \frac{1}{z^5} \ldots$$  \hfill (14.123)

There follows that for $z_k = 0$

$$R(z_k) = b_1 = -\frac{1}{2}.$$  \hfill (14.124)

Example 13.8 Find $R(1)$ for the complex function

$$f(z) = \frac{e^z}{(z-1)^2}.$$  \hfill (14.125)

**Solution:** The function has isolated singular point at $z_k = 1$ and we need to expand the function

$$g(z) = e^z$$  \hfill (14.126)

about this point. The series expansion about this point is is given by

$$g(z) = e^z = e \left[1 + \frac{1}{1!} (z-1) + \frac{1}{2!} (z-1)^2 + \frac{1}{3!} (z-1)^3 + \frac{1}{4!} (z-1)^4 \ldots \right]$$  \hfill (14.127)

then

$$f(z) = \frac{e^z}{(z-1)^2}.$$  \hfill (14.128)

becomes

$$f(z) = \frac{e}{(z-1)^2} \left\{1 + \frac{1}{1!} (z-1) + \frac{1}{2!} (z-1)^2 + \frac{1}{3!} (z-1)^3 + \frac{1}{4!} (z-1)^4 \ldots \right\}$$

$$= \frac{e}{(z-1)^2} + \frac{1}{2!} e + \frac{1}{3!} (z-1) + \frac{1}{4!} (z-1)^2 \ldots$$

There follows that the residue for $z_k = 1$,

$$R(z_k) = b_1 = e.$$  \hfill (14.129)
Example 13.9 Find $R(\pi)$ for the complex function

$$f(z) = \frac{\sin z}{z - \pi} \quad (14.130)$$

**Solution:** The function has an isolated singular point at

$$z_k = \pi \quad (14.131)$$

and we need to expand the function $\sin(z)$ in series about this point

$$\sin(z) = -(z - \pi) + \frac{(z - \pi)^3}{3!} - \frac{(z - \pi)^5}{5!} + \frac{(z - \pi)^7}{7!} \ldots \quad (14.132)$$

so that we can write

$$f(z) = \frac{\sin z}{z - \pi} \quad (14.133)$$

as

$$f(z) = -1 + \frac{(z - \pi)^2}{3!} - \frac{(z - \pi)^4}{5!} + \frac{(z - \pi)^6}{7!} \ldots \quad (14.134)$$

There follows that for $z_k = \pi$

$$R(z_k) = b_1 = 0. \quad (14.135)$$

**Method 2:** If $f(z)$ has a simple pole (a pole of order 1) at $z_o$, then multiply $f(z)$ by $(z - z_o)$ and evaluate the result at $z = z_o$ (or take the limit as $z$ approaches $z_o$):

$$R(z_o) = \lim_{z \to z_o} [(z - z_o) f(z)] \quad (14.136)$$

Example 13.10 Find $R(-1)$ for the complex function

$$f(z) = \frac{z}{(z + 1)(z + 2)} \quad (14.137)$$

**Solution:** We note that $f(z)$ has a simple pole (a order of 1) at $z_o = -1$. Then we can use the second method discussed above to find the residue $z_o = -1$

$$R(-1) = \lim_{z \to -1} [(z + 1) f(z)] = \lim_{z \to -1} \left[ (z + 1) \frac{z}{(z + 1)(z + 2)} \right]$$

$$\Rightarrow R(-1) = \lim_{z \to -1} \left[ \frac{z}{(z + 2)} \right] = \frac{-1}{(-1 + 2)} = -1. \quad (14.138)$$

Example 13.11 Find $R(\pi)$ for the complex function

$$f(z) = \cot(z) \quad (14.139)$$
Solution: The cotangent function is given by

\[ f(z) = \cot(z) = \frac{\cos(z)}{\sin(z)}. \quad (14.140) \]

Let’s assume that this function has a simple pole at \( z_0 = \pi \) which is true as we will see shortly. Then using the second method the residue at \( z_0 = \pi \) can be expressed as

\[
R(\pi) = \lim_{z \to \pi} [(z - \pi) f(z)] \Rightarrow R(\pi) = \lim_{z \to \pi} \left( (z - \pi) \frac{\cos(z)}{\sin(z)} \right) \quad (14.141)
\]

As \( z \to \pi \), the limit becomes \( 0/0 \) and we can apply L’Hospital rule

\[
R(\pi) = \lim_{z \to \pi} [(z - \pi) f(z)] \Rightarrow R(\pi) = \lim_{z \to \pi} \left( (z - \pi) \frac{\cos(z)}{\sin(z)} \right)\\
\Rightarrow R(\pi) = \frac{\cos(\pi) - (\pi - \pi) \sin(\pi)}{\cos(\pi)} = 1 \quad (14.142)
\]

Now if we make a series expansion for the function \( f(z) = \cot(z) = \frac{\cos(z)}{\sin(z)}. \quad (14.143) \)

we find

\[
f(z) = \frac{1}{z - \pi} - \frac{z - \pi}{3} - \frac{1}{45} (z - \pi)^3 - \frac{2}{945} (z - \pi)^5 - \frac{6}{4725} (z - \pi)^7 + \ldots \quad (14.144)
\]

which indeed shows the function has a simple pole at \( z_0 = \pi \).

Method 3 (Generalization of method 2): To find the residue of \( f(z) \) at \( z_0 \) when \( f(z) \) has a pole of order \( n \) at \( z_0 \), compute

\[
R(z_0) = \left[ \frac{1}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} \{(z - z_0)^m f(z)\} \right]_{z=z_0}, \quad (14.145)
\]

where \( m \geq n \).
Example 13.12 Find $R(3)$ for

$$f(z) = \frac{ze^{zt}}{(z-3)^2}$$  \hspace{1cm} (14.146)

where $t$ is a parameter (possible complex).

Solution: We note that $f(z)$ has a pole at $z_o = 3$ of order $n = 2$. Thus we can apply the third method to find the residue at $z_o$ which we write as

$$R(3) = \left[ \frac{d}{dz} \left( \frac{ze^{zt}}{(z-3)^2} \right) \right]_{z=3}$$

$$\Rightarrow \, R(3) = \left[ (1+zt)e^{zt} \right]_{z=3} = (1+3t)e^{3t}. \hspace{1cm} (14.147)$$

Example 13.13 Find $R(-2)$ for the complex function

$$f(z) = \frac{e^{2z}}{z(z+2)^3}$$  \hspace{1cm} (14.148)

Solution: The function has a pole at $z_o = -2$ of order $n = 3$. Using the third method for $m = n = 3$, we have

$$R(-2) = \frac{1}{(3-1)!} \left\{ \frac{d^{3-1}}{dz^{3-1}} \left( z(z+2)^3 \frac{e^{2z}}{z} \right) \right\} \bigg|_{z=-2} = \frac{1}{2!} \left[ 4e^{2z} - \frac{2e^{2z}}{z} + \frac{2e^{2z}}{z^2} + \frac{2e^{2z}}{z^3} \right] \bigg|_{z=-2}$$

$$= \frac{2e^{2z}}{z} - \frac{2e^{2z}}{z^2} + \frac{e^{2z}}{z^3} \bigg|_{z=-2} = \left( -1 - \frac{1}{2} - \frac{1}{8} \right) e^4 \Rightarrow R(-2) = -\frac{13}{8} e^{-4}. \hspace{1cm} (14.149)$$

14.9 Applications of the Residue Theorem

The Residue Theorem: we recall that if \( \{z_1, z_2, \ldots, z_n\} \) are singular points of \( f(z) \) inside a closed curve defined by \( C \), the residue theorem states that

$$\oint_C f(z) \, dz = 2\pi i \sum_{k=1}^n R(z_k),$$  \hspace{1cm} (14.150)

where \( R(z_k) \) is the residue of the function \( f(z) \) at \( z = z_k \). We are going to use this theorem and the methods of finding residues to evaluate several different types of definite integrals. The methods are best shown by examples.

Example 13.14 Evaluate the real integral

$$I = \int_0^\infty \frac{dx}{x^4 + 1}$$  \hspace{1cm} (14.151)
**Solution:** Consider the a semicircle of radius \( R \) on the upper half complex plane and the complex function

\[
f(z) = \frac{1}{z^4 + 1}.
\]  

(14.152)

Now let’s integrate the complex function over the semicircular contour in the upper half plane in the counterclockwise direction as shown in the figure below which we write as

\[
\oint \frac{dz}{z^4 + 1} = \int_{-R}^{R} \frac{dx}{x^4 + 1} + \int_{C} \frac{dz}{z^4 + 1}
\]  

(14.153)

For the curved part, we have

\[
z = R \exp[i\theta] \Rightarrow dz = iR \exp[i\theta] d\theta,
\]  

(14.154)

where \( 0 \leq \theta \leq \pi \), then

\[
\oint \frac{dz}{z^4 + 1} = \int_{-R}^{R} \frac{dx}{x^4 + 1} + \int_{0}^{\pi} \frac{iR \exp[i\theta] d\theta}{R^4 \exp[4i\theta] + 1}.
\]  

(14.155)

Now if we let \( R \to \infty \), we can see that the integral over the curved part (the second integral) approaches to zero and we find that

\[
\oint \frac{dz}{z^4 + 1} = \int_{-\infty}^{\infty} \frac{dx}{x^4 + 1}.
\]  

(14.156)
14.9. APPLICATIONS OF THE RESIDUE THEOREM

since \( x^4 + 1 \) is an even function we can write

\[
\int_{-\infty}^{\infty} \frac{dx}{x^4 + 1} = 2 \int_{0}^{\infty} \frac{dx}{x^4 + 1} \Rightarrow \int_{0}^{\infty} \frac{dx}{x^4 + 1} = \frac{1}{2} \int \frac{dz}{z^4 + 1}. \tag{14.157}
\]

The integral of the complex function

\[ f(z) = \frac{1}{z^4 + 1} \tag{14.158} \]

on the closed curve defined by this semicircle of infinite radius, using the Residue theorem, can be expressed as

\[
\oint f(z) \, dz = 2\pi i \sum_{k=1}^{n} R(z_k). \tag{14.159}
\]

In order to find the residues we first need to get the poles. Noting that

\[ z^4 + 1 = 0 \Rightarrow z = (-1)^{1/4} = \left[ e^{i(\pi/n(2\pi))} \right]^{1/4}, \tag{14.160} \]

where \( n = 0, 1, 2, 3 \) [Phys 3150 Roots of a Complex number], the poles for the function \( f(z) \) are found to be

\[
\begin{align*}
  z_0 &= e^{i\frac{\pi}{4}} \Rightarrow z_0 = \frac{1}{\sqrt{2}} (1 + i), \\
  z_1 &= e^{i\frac{3\pi}{4}} \Rightarrow z_1 = \frac{1}{\sqrt{2}} (-1 + i), \\
  z_2 &= e^{i\frac{5\pi}{4}} \Rightarrow z_2 = -\frac{1}{\sqrt{2}} (1 + i), \\
  z_3 &= e^{i\frac{7\pi}{4}} \Rightarrow z_3 = \frac{1}{\sqrt{2}} (1 - i). \tag{14.161}
\end{align*}
\]

Remember, we are integrating on the upper half of the complex plane. Hence, we are interested in the poles with a phase within \( 0 \leq \theta \leq \pi \).

These poles as shown in Fig. ?? are

\[
\begin{align*}
  z_0 &= e^{i\frac{\pi}{4}} = \frac{\sqrt{2}}{2} (1 + i), \\
  z_1 &= e^{i\frac{3\pi}{4}} = \frac{\sqrt{2}}{2} (-1 + i). \tag{14.162}
\end{align*}
\]

We now determine the residue of these poles. We recall that the residue of a pole, \( z = z_k \) of order \( n \) is given by

\[
R(z_k) = \left[ \frac{1}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} \{(z - z_k)^m f(z)\} \right]_{z=z_k} \tag{14.163}
\]

where \( m \geq n \). We note that the complex function \( f(z) \) can then be expressed as

\[ f(z) = \frac{1}{z^4 + 1} = \frac{1}{(z - z_0)(z - z_1)(z - z_2)(z - z_3)} \tag{14.164} \]
This shows that the two poles
\[ z_0 = e^{i \frac{\pi}{4}} = \frac{1}{\sqrt{2}} (1 + i), \quad z_1 = e^{i \frac{3\pi}{4}} = \frac{1}{\sqrt{2}} (-1 + i), \quad (14.165) \]
are simple poles (of order one) and we can use the second method to find the residues at these poles
\[ R(z_k) = \lim_{z \to z_k} [(z - z_k) f(z)]. \quad (14.166) \]
For \( z_0 \) one can write
\[ R(z_0) = \lim_{z \to z_0} \left[ \frac{1}{(z - z_0)(z - z_1)(z - z_2)} \right] = 1 \frac{1}{(z_0 - z_1)(z_0 - z_2)(z_0 - z_3)}, \quad (14.167) \]
and using the values for \( z_0, z_1, z_2, \) and \( z_3, \) in Eq. (14.161), we find
\[ z_0 - z_1 = \sqrt{2}, \quad z_0 - z_2 = \sqrt{2} (1 + i), \quad z_0 - z_3 = i \sqrt{2} \quad (14.168) \]
that give the residue
\[ R(z_0) = R \left( e^{i \frac{\pi}{4}} \right) = \frac{1}{\sqrt{2} (1 + i)(i \sqrt{2})} = -\frac{1}{2 \sqrt{2} (1 - i)}. \quad (14.169) \]
Similarly, the residue at the second pole \( z_1 \) which is given by
\[ R(z_1) = \lim_{z \to z_1} \left[ \frac{1}{(z - z_1)(z - z_0)(z - z_2)} \right] \Rightarrow R(z_0) = \frac{1}{(z_0 - z_1)(z_0 - z_2)(z_0 - z_3)}, \quad (14.170) \]
using Eq. (14.161),
\[ z_1 - z_0 = -\sqrt{2}, \quad z_1 - z_2 = i \sqrt{2}, \quad z_1 - z_3 = \sqrt{2} (-1 + i) \quad (14.171) \]
we find
\[ R(z_1) = R \left( e^{i \frac{3\pi}{4}} \right) = \frac{1}{-\sqrt{2} i \sqrt{2} \sqrt{2} (-1 + i)} = \frac{1}{2 \sqrt{2} (i + 1)}. \quad (14.172) \]
Now using the residue theorem
\[ \oint f(z) \, dz = 2\pi i \sum_{k=1}^{n} R(z_k) \quad (14.173) \]
we can see that
\[ \oint \frac{dz}{z^4 + 1} = 2\pi i \left[ -\frac{1}{2\sqrt{2}} \left( 1 - i \right) + \frac{1}{2\sqrt{2}} \left( i + 1 \right) \right] = \pi i \left[ \frac{-i - 1 + 1 - i}{2\sqrt{2}} \right] \]
\[ \Rightarrow \oint \frac{dz}{z^4 + 1} = \pi i \left[ \frac{-2i}{2\sqrt{2}} \right]. \]

Therefore
\[ \int_{0}^{\infty} \frac{dx}{x^4 + 1} = \frac{1}{2} \oint \frac{dz}{z^4 + 1} \Rightarrow \int_{0}^{\infty} \frac{dx}{x^4 + 1} = \frac{\pi}{2\sqrt{2}}. \]  

**Example 13.15** Evaluate the real integral
\[ I = \int_{0}^{2\pi} \frac{\cos(3\theta)}{5 - 4\cos(\theta)} d\theta \]  

**Solution:** Let’s consider a unit circle on a complex plane shown in Fig.14.7
Using Euler’s formula one can express
\[ \cos(k\theta) = \frac{1}{2} \left( e^{i\theta} + e^{-i\theta} \right)^k = \frac{1}{2} \left( e^{i\theta} + e^{-i\theta} \right), \]
\[ \sin(k\theta) = \frac{1}{2i} \left( e^{i\theta} - e^{-i\theta} \right)^k = \frac{1}{2i} \left( e^{i\theta} - e^{-i\theta} \right). \]

We note that for a complex number \( z \) on a unit circle shown in Fig.14.7, we have
\[ z = e^{i\theta}, \]
so that
\[ \cos(\theta) = \frac{1}{2} \left( e^{i\theta} + e^{-i\theta} \right) = \frac{1}{2} \left( z + \frac{1}{z} \right) = \frac{1}{2} \left( \frac{z^2 + 1}{z} \right), \]

and
\[ \cos(3\theta) = \frac{1}{2} \left( e^{i\theta} + e^{-i\theta} \right)^3 = \frac{1}{2} \left( z^3 + \frac{1}{z^3} \right) = \frac{1}{2} \left( \frac{z^6 + 1}{z^3} \right). \]
Then the function
\[ f(\theta) = \frac{\cos(3\theta)}{5 - 4\cos(\theta)} \]
can be rewritten as
\[ f(\theta) = \frac{\cos(3\theta)}{5 - 4\cos(\theta)} = \frac{\frac{z^6}{2} + 1}{-2z^2(2z^2 - 5z + 2)} = f(z) \] (14.180)

For \( z = \exp(i\theta) \), noting that
\[ dz = i\exp(i\theta)\,d\theta \Rightarrow d\theta = -i\exp(-i\theta)\,dz \Rightarrow d\theta = -i\frac{dz}{z}, \] (14.181)

one can then express the given integral as
\[
I = \int_0^{2\pi} \frac{\cos(3\theta)}{5 - 4\cos(\theta)}\,d\theta = \oint_C \frac{i(z^6 + 1)}{2z^3(2z^2 - 5z + 2)}\,dz,
\] (14.182)

where \( C \) is the unit circle shown in Fig. 14.7 and
\[ f(z) = \frac{z^6 + 1}{2z^3(2z^2 - 5z + 2)} \] (14.183)
We can then use the Residue theorem
\[ \oint_C f(z) \, dz = 2\pi i \sum_{k=1}^{n} R(z_k) \] (14.184)
to find the integral. To this end, we first find out the residues inside the unit circle. Noting that
\[ 2z^2 - 5z + 2 = (2z - 1)(z - 2) \Rightarrow 2z^2 - 5z + 2 = 2(z - 1/2)(z - 2) \] (14.185)
we may write
\[ f(z) = \frac{i \left( z^6 + 1 \right)}{4(z - 0)^3 (z - 1/2)(z - 2)}. \] (14.186)
This expression shows that the function \( f(z) \) has poles \( z_1 = 0, z_2 = 1/2, z_3 = 2 \). However, only the first two poles \( z_1 = 0, z_2 = 1/2 \) are inside the unit circle. Thus we need only the residue of these two poles. The second pole, \( z_2 = 1/2 \) is a simple pole since the order is one and the

![Figure 14.8: The three poles for the function, \( f(z) \).](image)

residue can be expressed as
\[
R(1/2) = \lim_{z \to 1/2} \left[ (z - 1/2) f(z) \right] = \lim_{z \to 1/2} \left[ \frac{i \left( z^6 + 1 \right)}{4z^3 (z - 2)} \right] = \frac{i \left( \frac{1}{64} + 1 \right)}{\frac{1}{2} \left( \frac{1}{2} - 2 \right)}
\]
\[
= \frac{65}{64} \left( -\frac{4}{3} \right) \Rightarrow R(1/2) = -i \frac{65}{48}. \] (14.187)
The first pole, \( z_0 = 0 \), is not a simple pole. It has order \( n = 3 \). Thus we use the third method to find the residue for this pole

\[
R(z_0) = \left[ \frac{1}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} \left( \frac{1}{z-z_0} \right) \right]_{z=z_0}, \tag{14.188}
\]

where \( m \geq n \). For \( m = 3 \) we may write

\[
R(0) = \frac{1}{(3-1)!} \frac{d^3}{dz^3} \left( \frac{z^3 i (z^6 + 1)}{2 z^3 (2 z^2 - 5 z + 2)} \right) \bigg|_{z=0} = \frac{1}{2} \frac{d^2}{dz^2} \left( \frac{i (z^6 + 1)}{2 (2 z^2 - 5 z + 2)} \right) \bigg|_{z=0} \tag{14.189}
\]

\[
R(0) = \frac{1}{2} \left[ - \frac{6 i z^5 (4 z - 5)}{(2 z^2 - 5 z + 2)^2} + \frac{15 i z^4}{2 z^2 - 5 z + 2} + \frac{i (1 + z^6) (4 z - 5)^2}{(2 z^2 - 5 z + 2)^3} - \frac{2 i (1 + z^6)}{(2 z^2 - 5 z + 2)^2} \right]_{z=0} \]

\[
R(0) = \frac{i}{2} \left( \frac{(-5)^2}{2^3} - \frac{2}{2^2} \right) = \frac{i}{2} \left( \frac{25}{8} - \frac{1}{2} \right) = \frac{21}{16} \tag{14.190}
\]

Therefore, the integral gives

\[
\oint f(z) \, dz = 2 \pi i \sum_{k=1}^{n} R(z_k) = 2 \pi i \left( -i \frac{65}{48} + i \frac{21}{16} \right) = \frac{\pi}{12}
\]

\[
\Rightarrow \int_{0}^{2\pi} \frac{\cos(3\theta)}{5 - 4 \cos(\theta)} \, d\theta = \oint f(z) \, dz = \frac{\pi}{12} \tag{14.191}
\]

**Mathematica result:**

\[
\text{In[1]} = \int_{0}^{2\pi} \frac{\text{Cos}[3 \theta]}{5 - 4 \text{Cos}[\theta]} \, d\theta
\]

\[
\text{Out[1]} = \frac{\pi}{12}
\]

**Standard Methods of Integration using Contour Integrals: Method I:** A real integral over the entire \( x \)-axis

**Method 2:** A real integral involving sines and cosines from 0 to \( 2\pi \)
14.10 The Kramers-Kronig Relations

The Residue Theorem: we recall that

\[ \int f(z) \, dz = 2\pi i R(z_k) \]  \hspace{1cm} (14.192)

where \( R(z_k) \) is the residue of the function \( f(z) \) at \( z = z_k \). If there are more than one residue, the residue theorem states that

\[ \int f(z) \, dz = 2\pi i \sum_{k=1}^{n} R(z_k) \]  \hspace{1cm} (14.193)

Here the poles \( z_k \) must be inside the closed curve over which the contour integral is being carried on. What if there are poles on the curve. Which means the function \( f(z) \) is not analytic at \( z_k \) that is/are located on the curve. For example,
let’s say we are interested to evaluate the real integral

\[ I_1 = \int_{-\infty}^{\infty} \frac{dx}{x^2 - 1}. \]  

(14.194)

To solve this integral we need to consider the closed curve shown in the figure below and evaluate the contour integral

\[ I_2 = \oint f(z) \, dz, \]  

(14.195)

where

\[ f(z) = \frac{1}{z^2 - 1}. \]  

(14.196)

But we know that \( f(z) \) has simple poles at \( z'_1 = -1 \) and \( z'_2 = 1 \) which are on the contour. So how can we evaluate \( I_1 \) by directly using the residue theorem. We can evaluate such kind of integral using two different approaches. We need to either put the poles inside or outside the contour by modifying the semicircular contour if we are going to use the residue theorem directly. As we will see either of these approach lead to the same result which result in a modification of the residue theorem.

14.11 The revised Residue Theorem

It states that

\[ \oint f(z) \, dz = 2\pi i \left[ \sum_{k=1}^{n} R(z_k) + \frac{1}{2} \sum_{k=1}^{n} R(z'_k) \right]. \]  

(14.197)
where $z_k$ is the pole inside the contour and $z'_k$ is a simple pole on the contour.

**Proof:** By creating a detour around the pole on the curve one can put the pole either inside or outside the contour as shown in Fig. ?? Suppose we consider the contour that put the pole, $z'_k$ inside. We assumed there are no other poles inside this contour except $z'_k$ so that

$$\int f(z) \, dz = 2\pi i R (z'_k). \quad (14.198)$$

For the modified contour, we can write

$$\oint f(z) \, dz = \int_{-R}^{z_k - \epsilon} f(z) \, dz + \int_{C_1} f(z) \, dz + \int_{z_k + \epsilon}^{R} f(z) \, dz + \int_{C_R} f(z) \, dz \quad (14.199)$$

where $C'$ is the curved part of the contour with radius $R$ in Fig. ?? Noting
that
\[
\int_{C_s} f(z) \, dz = f(z) = \sum_{n=0}^{\infty} a_n \int_{C_s} (z - z_k^{'})^n \, dz + \sum_{n=1}^{\infty} \int_{C_s} \frac{b_n}{(z - z_k^{'})^n} \, dz
\]
\[
= \sum_{n=0}^{\infty} a_n \int_{\pi}^{2\pi} e^n e^{i\theta} i\theta e^{i\theta} \, d\theta + \sum_{n=1}^{\infty} \int_{\pi}^{2\pi} \frac{b_n}{e^n e^{i\theta} i\theta e^{i\theta}} \, d\theta
\]
\[
\Rightarrow \lim_{\varepsilon \to 0} \int_{C_s} f(z) \, dz = ib_1 \int_{\pi}^{2\pi} i\pi b_1 = i\pi R(z_k^{'}) \quad (14.200)
\]

one can write
\[
\lim_{\varepsilon \to 0} \int_{-e}^{e-k} f(z) \, dz + \lim_{e \to 0} \int_{C_s} f(z) \, dz + \lim_{e \to 0} \int_{e-k}^{e-k+\varepsilon} f(z) \, dz + \lim_{e \to 0} \int_{C_R} f(z) \, dz
\]
\[
= \int_{-e}^{e-k} f(z) \, dz + i\pi R(z_k^{'}) + \int_{e-k}^{R} f(z) \, dz + \int_{C_R} f(z) \, dz = 2\pi i R(z_k^{'})
\]
\[
= \int_{-e}^{e-k} f(z) \, dz + \int_{e-k}^{R} f(z) \, dz + \int_{C_R} f(z) \, dz = 2\pi i R(z_k^{'}) - i\pi R(z_k^{'})
\]
\[
\Rightarrow \oint f(z) \, dz = \pi i R(z_k^{'}) \quad (14.201)
\]

**Example 13.15** Evaluate the integrals

\[
I_1 = \int_0^\infty \sin(x) \, \frac{dx}{x}, \quad I_2 = \int_{-\infty}^\infty \cos(x) \, \frac{dx}{x} \quad (14.202)
\]

**Solution:** Consider the function

\[
f(z) = \frac{e^{iz}}{z} \quad (14.203)
\]

For a semi-circular closed curve of radius, $R$, with diameter on the x-axis, this function has pole at $z_k^{'0} = 0$ which is exactly on the curve. Therefore, using the applying the revised residue theorem, we may write

\[
\oint \frac{e^{iz}}{z} \, dz = 2\pi i \left[ \sum_{k=1}^{n} R(z_k) + \frac{1}{2} \sum_{k=1}^{n} R(z_k^{'}) \right] \quad (14.204)
\]

\[
\oint \frac{e^{iz}}{z} \, dz = \pi i R(z_k^{'0}) = 0 \quad (14.205)
\]

This pole is first order pole and we can find the residue using

\[
R(z_k^{'0}) = \lim_{z \to 0} \left[ (z - z_k^{'0}) f(z) \right] \quad (14.206)
\]

which gives

\[
R(0) = \lim_{z \to 0} \left[ \frac{e^{iz}}{z} \right] = 1 \quad (14.207)
\]
Therefore the integral becomes

\[ \int \frac{e^{iz}}{z} \, dz = \pi i \]  
(14.208)

Now we note that in the limit \( R \to \infty \), we have

\[ \int e^{iz} \, dz = \int e^{ix} \, dx + \int C \frac{e^{iz}}{z} \, dz = \int \cos(x) \, dx + i \int \sin(x) \, dx \]

we may write

\[ \int \frac{\cos(x)}{x} \, dx + i \int \frac{\sin(x)}{x} \, dx = \pi i \]  
(14.210)

which leads to

\[ \int \frac{\sin(x)}{x} \, dx = \pi, \int \frac{\cos(x)}{x} \, dx = 0 \]  
(14.211)
Since $\frac{\sin(x)}{x}$ is an even function, we find

$$\int_{0}^{\infty} \frac{\sin(x)}{x} \, dx = \frac{\pi}{2}.$$  \hspace{1cm} (14.212)

*Disperion (The Frequency Dependence of permittivity)* The electrons in a nonconducting medium are bound to specific molecules. These electrons oscillates about the equilibrium position with a small amplitude. The electrons experiencing this kind of motion can be modeled as a harmonic oscillator. Then the electron experiences a spring force $F_s$ given by

$$F_s = -kx.$$  

In addition to this force, the electron can also experience some damping force $F_d$ which is proportional to the speed of the electron

$$F_d = -\gamma m \frac{dx}{dt}.$$  

If the electron is exposed to an EM wave with frequency $\omega$ and polarized in the x-direction

$$\vec{E} = E_0 \cos(\omega t) \hat{x}$$

it will experience a driving force

$$F_d = qE = qE_0 \cos(\omega t)$$

Then using newton's second law, the net force acting on the electron can be written as

$$m \frac{d^2 x}{dt^2} = qE_0 \cos(\omega t) - \gamma \frac{dx}{dt} - kx$$

$$\Rightarrow m \frac{d^2 x}{dt^2} + \gamma \frac{dx}{dt} + kx = qE_0 \cos(\omega t)$$

$$\Rightarrow \frac{d^2 x}{dt^2} + \frac{\gamma}{m} \frac{dx}{dt} + \frac{k}{m} x = \frac{qE_0}{m} \cos(\omega t)$$

or

$$\frac{d^2 x}{dt^2} + \frac{\gamma}{m} \frac{dx}{dt} + \omega_0^2 x = \frac{qE_0}{m} \cos(\omega t)$$

where

$$\omega_0 = \sqrt{\frac{k}{m}}$$

is the natural frequency of the electron. In terms of the complex variables $\hat{x}$ and $\hat{E}$, we may write this equation as

$$\frac{d^2 \hat{x}}{dt^2} + \gamma \frac{d\hat{x}}{dt} + \omega_0^2 \hat{x} = \frac{qE_0}{m} \exp(-i\omega t)$$  \hspace{1cm} (14.213)
The Homogenous solution to this differential equation is given by

\[ \tilde{x}_H(t) = e^{-\gamma t/2}(A \cos(\beta t) + B \sin(\beta t)) . \]

Substituting a particular solution of the form

\[ \tilde{x}_p(t) = \tilde{x}_0 \exp(-i\omega t) \]

into the differential equation, we find

\[ -m\omega^2 \tilde{x}_0 - i\gamma \omega m \tilde{x}_0 + k \tilde{x}_0 = qE_0 \]

which gives

\[ -m\omega^2 \tilde{x}_0 - im\gamma \omega \tilde{x}_0 + m\omega_0^2 \tilde{x}_0 = qE_0 \]

\[ \Rightarrow \tilde{x}_0 = \frac{qE_0/m}{\omega_0^2 - \omega^2 - i\gamma \omega} \]  
(14.215)

and the particular solution becomes

\[ \tilde{x}_p(t) = \frac{qE_0/m}{\omega_0^2 - \omega^2 - i\gamma \omega} \exp(-i\omega t) \]

Therefore the complex displacement of the electron is given by

\[ \tilde{x}(t) = \tilde{x}_H(t) + \tilde{x}_p(t) \]

\[ \tilde{x}(t) = e^{-\gamma t/2}(A \cos(\beta t) + B \sin(\beta t)) \]

\[ + \frac{qE_0/m}{\omega_0^2 - \omega^2 - i\gamma \omega} \exp(-i\omega t) \]

But we are interested in the steady state of the electron which happen if we waited long enough (i.e. \( t \to \infty \)). Thus for steady state the complex displacement of the electron becomes

\[ \tilde{x}(t) \simeq \frac{qE_0/m}{\omega_0^2 - \omega^2 - i\gamma \omega} \exp(-i\omega t) \]

which shows at steady state the electron begins to oscillates with frequency of the EM field as we would expect. The complex dipole moment of the electron can then be expressed as

\[ \tilde{p}(t) = q \tilde{x}(t) = \frac{q^2 E_0/m}{\omega_0^2 - \omega^2 - i\gamma \omega} \exp(-i\omega t) \]

This can also be put in the form

\[ \tilde{p}(t) = \frac{q^2 E_0/m}{(\omega_0^2 - \omega^2 + i\gamma \omega)^2 + (\gamma \omega)^2} \exp(-i\omega t) \]

\[ \Rightarrow \tilde{p}(t) = |\tilde{p}_0| \exp[-(i\omega t - \varphi)] \]
where
\[
|\vec{p}_0| = \frac{q^2 E_0/m}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma \omega)^2}}
\]
and
\[
\varphi = \tan^{-1} \left( \frac{\gamma \omega}{\omega_0^2 - \omega^2} \right).
\]
This means that \( p \) is out of phase by \( \varphi \) with respect to the electric field \( \vec{E} \).
Lagging behind by \( \varphi \) which is very small when \( \omega \ll \omega_0 \) and rises to \( \pi \) when \( \omega \gg \omega_0 \). If there are \( N \) molecules per unit volume and \( f_j \) electrons per molecule. If these electrons in the \( j^{th} \) molecule is oscillating with the natural frequency \( \omega_j \) and damped by \( \gamma_j \), then the total dipole moment of the electrons per unit volume (the polarization \( \vec{P} \)) can be expressed as
\[
\vec{P} = \frac{Nq^2}{m} \sum_{j=1}^{\infty} \frac{f_j}{\omega_j^2 - \omega^2 - i\gamma_j \omega} E_0 \exp(-i\omega t)
\]
This can also be put in the form
\[
\vec{P} = \epsilon_0 \chi_e \vec{E}
\]
where
\[
\chi_e = \frac{Nq^2}{\epsilon_0 m} \sum_{j=1}^{\infty} \frac{f_j}{\omega_j^2 - \omega^2 - i\gamma_j \omega}.
\]
Recalling that the complex permittivity
\[
\tilde{\epsilon} = \epsilon_0 \left( 1 + \chi_e \right)
\]
and the complex dielectric constant can be expressed as
\[
\tilde{\epsilon}_r = \frac{\tilde{\epsilon}}{\epsilon_0} = 1 + \chi_e = 1 + \frac{Nq^2}{\epsilon_0 m} \sum_{j=1}^{\infty} \frac{f_j}{\omega_j^2 - \omega^2 - i\gamma_j \omega} \quad (14.216)
\]
Eq. (14.216) shows the dielectric constant (the permittivity of the medium) depends on the frequency of the EM wave-the medium is dispersive. Still wave equation is satisfied by the electric and magnetic fields. For a dispersive medium it is given by
\[
\nabla^2 \tilde{E} = \tilde{\mu}_0 \frac{\partial^2 \tilde{E}}{\partial t^2}
\]
with a plane wave
\[
\tilde{E} = \tilde{E}_0 \exp \left[ i (kz - \omega t) \right].
\]
where the wave number $\tilde{k}$ is complex and is given by

$$\frac{\omega}{k} = \frac{1}{\sqrt{\epsilon_0\mu_0}} \Rightarrow \tilde{k} = \sqrt{\epsilon_0\mu_0}\omega$$

If we express it as

$$\tilde{k} = k_{\text{Re}} + ik_{\text{Im}}$$

The electric field may be put in the form

$$\tilde{E} = \tilde{E}_0 \exp(-k_{\text{Im}}z) \exp[i (k_{\text{Re}}z - \omega t)]$$

and the intensity which is proportional to the square of the amplitude of the electric field becomes

$$I \propto |\tilde{E}_0|^2 \exp(-2k_{\text{Im}}z)$$

which shows a damping in the field amplitude due to the absorption by the medium. For that reason the absorption coefficient of the medium is given by

$$\alpha = 2k_{\text{Im}}.$$ 

The wave velocity is given by

$$v = \frac{\omega}{k_{\text{Re}}}$$

and the refractive index of the medium

$$n = \frac{c}{v} = \frac{\epsilon_0\mu_0}{\epsilon_0\mu_0} \frac{\omega}{c}$$

The complex wave number

$$\tilde{k} = \sqrt{\epsilon_0\mu_0}\omega = \sqrt{\frac{\epsilon}{\epsilon_0}\mu_0}\omega = \sqrt{\epsilon_r\frac{\omega}{c}}$$

$$\Rightarrow \tilde{k} = \frac{\omega}{c} \left[ 1 + \frac{Nq^2}{\epsilon_0me} \sum_{j=1}^{\infty} \frac{f_j}{\omega_j^2 - \omega^2 - i\gamma_j \omega} \right]^{1/2}$$

For gases the second terms is very small and we can make a binomial expansion for the square root which gives

$$\tilde{k} \simeq \frac{\omega}{c} \left( 1 + \frac{Nq^2}{2\epsilon_0me} \sum_{j=1}^{\infty} \frac{f_j}{\omega_j^2 - \omega^2 - i\gamma_j \omega} \right).$$

Then we can find the real and imaginary part of the wave number to be

$$k_{\text{Re}} \simeq \frac{\omega}{c} \left( 1 + \frac{Nq^2}{2\epsilon_0me} \sum_{j=1}^{\infty} \frac{f_j}{\omega_j^2 - \omega^2 - (\gamma_j \omega)^2} \right)$$

$$k_{\text{Im}} \simeq \frac{Nq^2\omega^2}{2\epsilon_0me} \sum_{j=1}^{\infty} \frac{f_j \gamma_j}{(\omega_j^2 - \omega^2)^2 + (\gamma_j \omega)^2}$$
so that the absorption coefficient and the refractive index of the medium can be expressed as

\[
\alpha = \frac{N q^2 \omega^2}{\epsilon_0 m} \sum_{j=1} f_j \gamma_j \frac{1}{(\omega_j^2 - \omega^2)^2 + (\gamma_j \omega)^2}
\]

\[
n = 1 + \frac{N q^2}{2 \epsilon_0 m} \sum_{j=1} \frac{f_j (\omega_j^2 - \omega^2)}{(\omega_j^2 - \omega^2)^2 + (\gamma_j \omega)^2}
\]

For just one particular molecule \( (j^\text{th} \text{ molecule}) \) we may write that

\[
\alpha = \frac{N q^2 \omega^2}{\epsilon_0 m} \frac{f_j \gamma_j}{(\omega_j^2 - \omega^2)^2 + (\gamma_j \omega)^2}
\]

\[
n = 1 + \frac{N q^2}{2 \epsilon_0 m} \frac{f_j (\omega_j^2 - \omega^2)}{(\omega_j^2 - \omega^2)^2 + (\gamma_j \omega)^2}
\]

Introducing the transformation defined by the dimensionless variables

\[
\lambda = \frac{\omega}{\omega_j}, \eta = \frac{\gamma_j}{\omega_j}
\]

we can write

\[
\alpha = \frac{N q^2 \omega_j^3}{\epsilon_0 m \omega_j} \frac{\lambda^2 \eta}{(1 - \lambda^2)^2 + (\eta \lambda)^2}
\]

\[
n = 1 + \frac{N q^2 \omega_j^3}{2 \epsilon_0 m \omega_j^4} \frac{f_j (1 - \lambda^2)}{(1 - \lambda^2)^2 + (\eta \lambda)^2}
\]

\[
\alpha = \frac{N q^2}{\epsilon_0 m \omega_j} \frac{\lambda^2 \eta}{(1 - \lambda^2)^2 + (\eta \lambda)^2}
\]

\[
n = 1 + \frac{N q^2 f_j}{2 \epsilon_0 m \omega_j^2} \frac{(1 - \lambda^2)}{(1 - \lambda^2)^2 + (\eta \lambda)^2}
\]

\[
\alpha = \Gamma_1 \frac{\lambda^2 \eta}{(1 - \lambda^2)^2 + (\eta \lambda)^2}
\]

\[
n = 1 + \Gamma_2 \frac{(1 - \lambda^2)}{(1 - \lambda^2)^2 + (\eta \lambda)^2}
\]

where

\[
\Gamma_1 = \frac{N q^2}{\epsilon_0 m \omega_j}, \Gamma_2 = \frac{N q^2 f_j}{2 \epsilon_0 m \omega_j^2} = \frac{\Gamma_1 f_j}{2 \omega_j}
\]

The plot for such particular molecule the absorption coefficient is shown in figure below. It shows that as the damping constant increases the range of the spectrum that would be absorbed by the medium increases with a peak
absorption at resonance when the frequency of the EM wave is the same as the natural frequency of the electrons (i.e. $\lambda = 1 \Rightarrow \omega = \omega_j$).

On the other hand the refractive index shown in the figure below shows unusual behavior near the resonant region. In this region the refractive index decreases in contrary what we know in optics. The range of the spectrum where this unusual behavior is observed increases with increase in damping effect.

Comparison of the absorption coefficient and the refractive index is shown in the figure below. The curves shown by the dotted lines are for the absorption and the solid line are for the refractive index. In the immediate neighborhood of a resonance the index of refraction drops sharply. The material may be practically opaque in this frequency range since it coincides with the region of maximum absorption. This is because of the electron are forced to oscillate with their favorite frequency and the amplitude of oscillation is maximum and correspondingly high amount of energy is dissipated by the damping mechanism. Because this behavior is atypical, it is called anomalous dispersion (the region $\omega_1 < \omega < \omega_2$).

Refractive index increases for $\omega < \omega_1$ and $\omega > \omega_2$ which is consistent with our experience from optics (dispersion). The most familiar example of dispersion is probably a rainbow, in which dispersion causes the spatial separation of a white light into components of different wavelengths (different colors).

Refractive index decreases for $\omega_1 < \omega < \omega_2$, in the immediate neighborhood of a resonance the index of refraction drops sharply. The material may be practically opaque in this frequency range since it coincides with the region of maximum absorption. This is because of the electron are forced to oscillate with their favorite frequency and the amplitude of oscillation is maximum and
correspondingly high amount of energy is dissipated by the damping mechanism. Because this behavior is atypical, it is called anomalous dispersion (the region $\omega_1 < \omega < \omega_2$).

Refractive index less than one for $\omega > \omega_2$ which means the wave speed wave exceeds $c$. This is not an alarm since energy does not travel at the wave velocity but rather at a group velocity. Moreover we considered only one molecule.

Far away from resonance ($\omega >> \omega_j$): For the case where we are far away from resonance we can ignore the damping term and we can write $(\omega_j^2 - \omega^2)^2 + (\gamma_j \omega)^2 \approx (\omega_j^2 - \omega^2)^2$. Then index of refraction can be expressed as

$$n = 1 + \frac{Nq^2}{2\epsilon_0 m} \sum_{j=1}^{\infty} \frac{f_j}{(\omega_j^2 - \omega^2)}$$

For transparent materials, the nearest significant resonance typically lie in the ultraviolet, then $\omega < \omega_j$. Taking this into account and noting that we are very far from resonance we can make the approximation

$$\frac{1}{\omega_j^2 - \omega^2} = \frac{1}{\omega_j^2} \frac{1}{1 - \omega^2} \approx \frac{1}{\omega_j^2} \left(1 + \frac{\omega^2}{\omega_j^2}\right)$$
If we express the frequency in terms of the wave length in a vacuum

\[ \omega = \frac{2\pi c}{\lambda} \]

we can write

\[ n = 1 + \left( \frac{Nq^2}{2\epsilon_0 m} \sum_{j=1}^{\infty} \frac{f_j}{\omega_j^2} \right) + \frac{1}{\lambda^2} \left( \frac{2\pi^2 c^2 Nq^2}{\epsilon_0 m} \sum_{j=1}^{\infty} \frac{f_j}{\omega_j^4} \right) \]

or

\[ n = 1 + A \left( 1 + \frac{B}{\lambda^2} \right) \]

This is known as Cauchy’s Formula: the constant \( A \) is called the coefficient of Refraction and \( B \) is called the coefficient of Dispersion. It applies in most gases, in the optical region.

The Kramers-Kronig Relations: Consider the complex function

\[ f(\omega) = u(\omega) + iv(\omega) \quad (14.217) \]
which is a complex function $\omega$, where $u(\omega)$ and $v(\omega)$ are real. Suppose this function is analytic in the upper half-plane of $\omega$ and it vanishes faster than $1/|\omega|$ as $|\omega| \to \infty$. The Kramers-kronig relations are given by

$$f(\omega_0) = \frac{1}{\pi i} P \int_{\infty}^{-\infty} \frac{f(\omega)}{\omega - \omega_0} d\omega = \frac{1}{\pi} P \int_{\infty}^{-\infty} \frac{u(\omega) + iv(\omega)}{\omega - \omega_0} d\omega$$

$$\Rightarrow u(\omega) + iv(\omega) = \frac{1}{\pi} P \int_{\infty}^{-\infty} \frac{v(\omega)}{\omega - \omega_0} d\omega - \frac{1}{\pi} P \int_{\infty}^{-\infty} \frac{u(\omega)}{\omega - \omega_0} d\omega$$

(14.218)

which leads to

$$u(\omega) = \frac{1}{\pi} P \int_{\infty}^{-\infty} \frac{v(\omega)}{\omega - \omega_0} d\omega, \quad v(\omega) = -\frac{1}{\pi} P \int_{\infty}^{-\infty} \frac{u(\omega)}{\omega - \omega_0} d\omega,$$

(14.219)

where $P$ indicates the Cauchy principal value which is given by

$$P \int_{\infty}^{-\infty} \frac{u(\omega)}{\omega - \omega_0} d\omega = \lim_{\varepsilon \to 0} \left[ \int_{-\infty}^{\infty} \frac{u(\omega)}{\omega - \omega_0} d\omega + \int_{\omega_0 + \varepsilon}^{\infty} \frac{u(\omega)}{\omega - \omega_0} d\omega \right]$$

(14.220)

**Dispersion:** The Kramers-Kronig Relations has a wide application in a dispersive medium (frequency dependent) for Electromagnetic waves. In such medium the propagation vector $k$ is a complex quantity which leads to a complex refractive index

$$n^2(\omega) = 1 + i \frac{4\pi \sigma}{\omega} \Rightarrow n^2(\omega) - 1 = \frac{4\pi \sigma}{\omega}$$

(14.221)

According to the Kramers-Kronig Relations, we can write

$$\text{Re} \left[ n^2(\omega) - 1 \right] = \frac{2}{\pi} P \int_{0}^{\infty} \frac{\omega \text{Im} \left[ n^2(\omega) - 1 \right]}{\omega^2 - \omega_0^2} d\omega$$

(14.222)

$$\text{Im} \left[ n^2(\omega) - 1 \right] = -\frac{2}{\pi} P \int_{0}^{\infty} \frac{\omega_0 \text{Re} \left[ n^2(\omega) - 1 \right]}{\omega^2 - \omega_0^2} d\omega$$

(14.223)

where $\omega_0$ is a general real angular frequency.
Chapter 15

Laplace transform

Integral Transforms: Frequently in physics we encounter pairs of functions related by an integral of the form

\[ F(p) = \int_a^b K(p, t) f(t) \, dt \quad (15.1) \]

The function \( F(p) \) is called the (integral) transform of \( f(t) \) by the kernel \( K(p, t) \). There are two types of integral transform that we will consider; the Laplace and Fourier transforms. Integral transform operation may also be described as mapping a function \( f(t) \) in \( t \)-space into another function \( F(p) \) in \( p \)-space. This interpretation takes on physical significance in the time-frequency relation of Fourier transforms.

Laplace Transforms: the Laplace transform \( F(p) \) or \( \mathcal{L} \) of a function \( f(t) \) is defined by

\[ F(p) = \mathcal{L}[f(t)] = \lim_{b \to \infty} \int_0^b e^{-pt} f(t) \, dt. \quad (15.2) \]

Note:

(i) The Laplace transform for the function \( f(t) \) exists even if the integral for this function,

\[ I = \int_0^\infty f(t) \, dt \quad (15.3) \]

does not exist.

(ii) For a function \( f(t) \) the Laplace transform to exist there must be a positive constant \( M \) such that

\[ |e^{-pt} f(t)| \leq M \quad (15.4) \]

for sufficiently large \( t, \ t > t_0 \).

\[ f(t) = e^{t^2}. \quad (15.5) \]
(iii) The Laplace transform fails to exist when the function has a strong singularity as \( t \to 0 \)

\[
f(t) = t^n \quad \text{for} \quad n \leq -1.
\] 

(15.6)

(iv) Laplace Transform is linear. That means if the Laplace transform for the two functions \( f(t) \) and \( g(t) \) exist, then we can write that

\[
\mathcal{L} [af(t) + dg(t)] = a\mathcal{L} [f(t)] + d\mathcal{L} [g(t)],
\]

where \( a \) and \( d \) are constants.

Example 14.1 Show that the Laplace Transform is linear

Solution: Consider the function

\[
h(t) = af(t) + dg(t) \tag{15.8}
\]

so that the Laplace transform of this function can be expressed as

\[
H(p) = \mathcal{L} [h(t)] = \lim_{b \to \infty} \int_0^b e^{-pt} [af(t) + dg(t)] \, dt
\]

\[
= \lim_{b \to \infty} \int_0^b e^{-pt} af(t) \, dt + \lim_{a \to \infty} \int_0^a e^{-pt} dg(t) \, dt = a \lim_{a \to \infty} \int_0^a e^{-pt} f(t) \, dt + \lim_{a \to \infty} \int_0^a e^{-pt} g(t) \, dt
\]

\[
\Rightarrow H(p) = \mathcal{L} [af(t) + bg(t)] = a\mathcal{L} [f(t)] + d\mathcal{L} [g(t)] \tag{15.9}
\]

Example 14.2 Find the Laplace transform of the following functions and specify the conditions [if there is any specific condition that must be satisfied to do the transform]

(a) 

\[
f(t) = 1 \tag{15.10}
\]

(b) 

\[
f(t) = e^{\omega t} \tag{15.11}
\]

c) 

\[
f_1(t) = \cosh(\omega t), \quad f_2(t) = \sinh(\omega t) \tag{15.12}
\]

d) 

\[
f(t) = \cos(\omega t), \quad f(t) = \sin(\omega t) \tag{15.13}
\]

Solution:
Recalling that
\[ F(p) = \mathcal{L} \left[ f(t) \right] = \int_0^\infty e^{-pt} f(t) \, dt. \]  
we find
\[ \mathcal{L}[1] = \int_0^\infty e^{-pt} \, dt = \frac{1}{p}. \]
for \( p > 0 \).

\( \textbf{(b)} \)

\[ F(p) = \mathcal{L} \left[ e^{\omega t} \right] = \int_0^\infty e^{-(p-\omega)t} \, dt = \frac{1}{p-\omega} \]
for \( p > \omega \). Replacing \( \omega \) by \(-\omega\), one can easily find
\[ F(p) = \mathcal{L} \left[ e^{-\omega t} \right] = \int_0^\infty e^{-(p+\omega)t} \, dt = \frac{1}{p+\omega}, \]
for \( p + \omega > 0 \).

\( \textbf{(c)} \)

Noting that
\[ \cosh(\omega t) = \frac{e^{\omega t} + e^{-\omega t}}{2}, \quad \sinh(\omega t) = \frac{e^{\omega t} - e^{-\omega t}}{2} \]
we may write
\[ F(p) = \mathcal{L} \left[ \cosh(\omega t) \right] = \mathcal{L} \left[ \frac{e^{\omega t} + e^{-\omega t}}{2} \right] \]
Recalling that Laplace transform is linear, one can write
\[ \mathcal{L} \left[ \cosh(\omega t) \right] = \frac{1}{2} \left( \mathcal{L} \left[ e^{\omega t} \right] + \mathcal{L} \left[ e^{-\omega t} \right] \right). \]
and applying the result in the previous example, we find
\[ \mathcal{L} \left[ \cosh(\omega t) \right] = \frac{1}{2} \left( \frac{1}{p-\omega} + \frac{1}{p+\omega} \right) = \frac{p}{p^2 - \omega^2}. \]
Similarly
\[ F(p) = \mathcal{L} \left[ \sinh(\omega t) \right] = \mathcal{L} \left[ \frac{e^{\omega t} - e^{-\omega t}}{2} \right] = \frac{1}{2} \left[ \mathcal{L} \left[ e^{\omega t} \right] - \mathcal{L} \left[ e^{-\omega t} \right] \right] \]
\[ \Rightarrow \mathcal{L} \left[ \sinh(\omega t) \right] = \frac{1}{2} \left( \frac{1}{p-\omega} - \frac{1}{p+\omega} \right) = \frac{\omega}{p^2 - \omega^2} \]

\( \textbf{(d)} \)

Recalling that
\[ \cos(\omega t) = \cosh(i\omega t), \quad \sin(\omega t) = -i \sinh(i\omega t) \]

CHAPTER 15. LAPLACE TRANSFORM

and applying the result above we can easily write that

\[ \mathcal{L} \left[ \cos (\omega t) \right] = \mathcal{L} \left[ \cosh (i\omega t) \right] = \frac{p}{p^2 - (i\omega)^2} = \frac{p}{p^2 + \omega^2} \]

\[ \mathcal{L} \left[ \sin (\omega t) \right] = \mathcal{L} \left[ -i \sinh (i\omega t) \right] = -i \mathcal{L} \left[ \sinh (i\omega t) \right] = \frac{-i (i\omega)}{p^2 - (i\omega)^2} \]

\[ \mathcal{L} \left[ \sin (\omega t) \right] = \frac{\omega}{p^2 + \omega^2} \]  \hspace{1cm} (15.24)

**Example 14.2** For the unit step function, also called the Heaviside function, defined by

\[ U(t - a) = \begin{cases} 0 & \text{for } t < a \\ 1 & \text{for } t > a \end{cases} \]  \hspace{1cm} (15.25)

Find

\[ \mathcal{L} \left[ U(t - a) \right]. \]  \hspace{1cm} (15.26)

**Solution:** The Laplace transform

\[ F(p) = \mathcal{L} [f(t)] = \int_0^\infty e^{-pt} f(t) \, dt. \]  \hspace{1cm} (15.27)

for the unit step function can be rewritten as

\[ \mathcal{L} [f(t)] = \int_0^a e^{-pt} f(t) \, dt + \int_a^\infty e^{-pt} f(t) \, dt = \int_a^\infty e^{-pt} \, dt = e^{-pa} \bigg|_a^\infty \]

\[ \Rightarrow \mathcal{L} [f(t)] = \frac{e^{-pa}}{p}. \]  \hspace{1cm} (15.28)

**Example 14.3** Find the Laplace transform for the first, second, and third derivative of the function, \( \frac{df(t)}{dt} \)

(a)

\[ \mathcal{L} \left[ f'(t) \right] = \mathcal{L} \left[ \frac{df(t)}{dt} \right], \]  \hspace{1cm} (15.29)

(b)

\[ \mathcal{L} \left[ f''(t) \right] = \mathcal{L} \left[ \frac{d^2 f(t)}{dt^2} \right], \]  \hspace{1cm} (15.30)

(c)

\[ \mathcal{L} \left[ f'''(t) \right] = \mathcal{L} \left[ \frac{d^3 f(t)}{dt^3} \right], \]  \hspace{1cm} (15.31)

and show that the general relation can be written as

\[ \mathcal{L} \left[ f^{(n)}(t) \right] = \mathcal{L} \left[ \frac{d^n f(t)}{dt^n} \right] = p^n \mathcal{L} [f(t)] - p^{n-1} f(0) - p^{n-2} f'(0) \]

\[ -p^{n-3} f''(0) - ... - f^{(n-1)}(0). \]  \hspace{1cm} (15.32)
Solution:

(a) Recalling that
\[ \mathcal{L} [f(t)] = \int_0^\infty e^{-pt} f(t) \, dt. \]  
(15.33)
for \( f'(t) \), one can write
\[ \mathcal{L} [f'(t)] = \int_0^\infty e^{-pt} \frac{df(t)}{dt} \, dt \]  
(15.34)
so that applying integration by parts we may write
\[ \mathcal{L} [f'(t)] = e^{-pt} f(t) \bigg|_0^\infty + p \int_0^\infty e^{-pt} f(t) \, dt = -f(0) + p \int_0^\infty e^{-pt} f(t) \, dt \]
\[ \Rightarrow \mathcal{L} [f'(t)] = p\mathcal{L} [f(t)] - f(0) \]

(b) Similarly for the second derivative
\[ \mathcal{L} [f''(t)] = \int_0^\infty e^{-pt} \left[ \frac{d^2 f(t)}{dt^2} \right] \, dt \]  
(15.35)
we have
\[ \mathcal{L} [f''(t)] = e^{-pt} \frac{df(t)}{dt} \bigg|_0^\infty + p \int_0^\infty e^{-pt} \frac{df(t)}{dt} \, dt \]
\[ = p \int_0^\infty e^{-pt} \frac{df(t)}{dt} \, dt - \frac{df(t)}{dt} \bigg|_{t=0} \]
\[ \Rightarrow \mathcal{L} [f''(t)] = p\mathcal{L} [f'(t)] - f'(0) \]  
(15.36)
so that using the result we found in part (a),
\[ \mathcal{L} [f'(t)] = p\mathcal{L} [f(t)] - f(0) \]  
(15.37)
we find
\[ \mathcal{L} [f''(t)] = p^2 \mathcal{L} [f(t)] - pf(0) - f'(0). \]  
(15.38)

(c) For the third derivative
\[ \mathcal{L} [f'''(t)] = \int_0^\infty e^{-pt} \frac{d^3 f(t)}{dt^3} \, dt = e^{-pt} \frac{d^2 f(t)}{dt^2} \bigg|_0^\infty + p \int_0^\infty e^{-pt} \frac{d^2 f(t)}{dt^2} \, dt \]
\[ \Rightarrow \mathcal{L} [f'''(t)] = p\mathcal{L} [f''(t)] - f''(0) \]  
(15.39)
and using the result in part (b), one finds
\[ \mathcal{L} [f'''(t)] = p^3 \mathcal{L} [f(t)] - p^2 f(0) - pf'(0) - f''(0) \]
Based on the result we derived in (a)-(c), one can then write
which leads to the general relation

\[ L \left[ f^{(n)} (t) \right] = L \left[ \frac{d^n f(t)}{dt^n} \right] = p^n L[f(t)] - p^{n-1} f(0) - p^{n-2} \frac{df(0)}{dt} - \ldots - \frac{d^{n-1} f(0)}{dt^{n-1}}. \] (15.40)

### 15.1 Inverse Laplace Transform

The inverse of a Laplace transform is represented by

\[ L^{-1} [F(p)] = f(t). \] (15.41)

In order to perform inverse Laplace transform it is important to know the
Laplace transform of some of the basic functions that we saw in the examples ear-
lier.

<table>
<thead>
<tr>
<th>( n )</th>
<th>( L \left[ \frac{d^n f(t)}{dt^n} \right] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( pL[f(t)] - f(0) )</td>
</tr>
<tr>
<td>2</td>
<td>( p^2L[f(t)] - pf(0) - p \frac{df(0)}{dt} )</td>
</tr>
<tr>
<td>3</td>
<td>( p^3L[f(t)] - p^2f(0) - p \frac{df(0)}{dt} - p \frac{d^2f(0)}{dt^2} )</td>
</tr>
<tr>
<td>4</td>
<td>( p^4L[f(t)] - p^3f(0) - p^2f(0) - p \frac{df(0)}{dt} - p \frac{d^2f(0)}{dt^2} - p \frac{d^3f(0)}{dt^3} )</td>
</tr>
<tr>
<td>5</td>
<td>( p^5L[f(t)] - p^4f(0) - p^3f(0) - p^2f(0) - p \frac{df(0)}{dt} - p \frac{d^2f(0)}{dt^2} - p \frac{d^3f(0)}{dt^3} - p \frac{d^4f(0)}{dt^4} )</td>
</tr>
</tbody>
</table>

---

**Example 14.4** Evaluate the following inverse Laplace transforms

(a) \( L^{-1} \left[ \frac{5}{p + 2} \right] \) (15.42)

(b) \( L^{-1} \left[ \frac{4p - 3}{p^2 + 4} \right] \) (15.43)

**Solution:**

(a) Noting that

\[ L^{-1} \left[ \frac{5}{p + 2} \right] = 5L^{-1} \left[ \frac{1}{p - (-2)} \right] \] (15.44)

and recalling

\[ L \left[ e^{\omega t} \right] = \frac{1}{p - \omega} \] (15.45)
we can write

\[ \mathcal{L}^{-1} \left[ \frac{5}{p+2} \right] = 5e^{-2t}. \]  

(15.46)

(b) Here also we want to write

\[ \mathcal{L}^{-1} \left[ \frac{4p-3}{p^2 + 4} \right] = \mathcal{L}^{-1} \left[ \frac{4p-3}{p^2 + 2^2} \right] = \mathcal{L}^{-1} \left[ \frac{4p}{p^2 + 2^2} - \frac{3}{p^2 + 2^2} \right] \]

\[ = \mathcal{L}^{-1} \left[ \frac{4p}{p^2 + 2^2} \right] - \mathcal{L}^{-1} \left[ \frac{3}{p^2 + 2^2} \right] \]

\[ \Rightarrow \mathcal{L}^{-1} \left[ \frac{4p-3}{p^2 + 4} \right] = 4\mathcal{L}^{-1} \left[ \frac{p}{p^2 + 2^2} \right] - \frac{3}{2} \mathcal{L}^{-1} \left[ \frac{2}{p^2 + 2^2} \right]. \]  

(15.47)

Recalling that

\[ \mathcal{L} \left[ \cos (\omega t) \right] = \frac{p}{p^2 + \omega^2}, \mathcal{L} \left[ \sin (\omega t) \right] = \frac{\omega}{p^2 + \omega^2} \]  

so that

\[ \mathcal{L}^{-1} \left[ \frac{4p-3}{p^2 + 4} \right] = 4\cos(2t) - \frac{3}{2} \sin(2t) \]  

(15.49)

**Mathematica Result:**

\[ \text{In}[1]:= \text{InverseLaplaceTransform}\left[\frac{4p-3}{p^2 + 4}, p, t\right] \]

\[ \text{Out}[1]= 4 \cos[2t] - 3 \cos[t] \sin[t] \]

**Example 14.5** Evaluate the integral

\[ f(t) = \int_0^\infty \frac{\sin(tx)}{x} \, dx \]  

(15.50)

applying Laplace and inverse Laplace transform

**Solution:** We first apply Laplace transform

\[ \mathcal{L} \left[ f(t) \right] = \int_0^\infty e^{-pt} \left[ \int_0^\infty \frac{\sin(tx)}{x} \, dx \right] \, dt \]  

(15.51)

which can also be expressed, by interchanging the order of the integration,
as

$$\mathcal{L} \left[ f(t) \right] = \int_0^\infty \frac{1}{x} \left[ \int_0^\infty \sin(tx) e^{-pt} dt \right] dx$$

$$= \int_0^\infty \frac{1}{x} \left[ \int_0^\infty \frac{e^{itx} - e^{-itx}}{2i} e^{-pt} dt \right] dx = \int_0^\infty \frac{1}{x} \left[ \int_0^\infty \frac{e^{-(p-ix)t} - e^{-(p+ix)t}}{2i} dt \right] dx$$

$$= \int_0^\infty \frac{1}{2ix} \left[ \frac{e^{-(p-ix)t}}{-p-ix} + \frac{e^{-(p+ix)t}}{p+ix} \right] dx$$

$$= \int_0^\infty \frac{1}{2ix} \left[ \frac{1}{p-ix} - \frac{1}{p+ix} \right] dx = \int_0^\infty \frac{1}{2ix} \left[ \frac{p+ix-p+ix}{p^2+x^2} \right] dx$$

$$\Rightarrow \mathcal{L} \left[ f(t) \right] = \int_0^\infty \frac{dx}{p^2+x^2} = \frac{\pi}{2p} \quad (15.52)$$

where we used the result we determined in using the Residue theorem.

**A short way:** We note that

$$\mathcal{L} \left[ f(t) \right] = \int_0^\infty \frac{\mathcal{L}[\sin(tx)]}{x} dx \quad (15.53)$$

and using the relation

$$\mathcal{L} \left[ \sin(\omega t) \right] = \frac{\omega}{p^2 + \omega^2} \quad (15.54)$$

we have

$$\mathcal{L} \left[ \sin(tx) \right] = \frac{x}{p^2 + x^2} \quad (15.55)$$

so that

$$\mathcal{L} \left[ f(t) \right] = \int_0^\infty \frac{\mathcal{L}[\sin(tx)]}{x} dx = \int_0^\infty \frac{dx}{p^2 + x^2} = \frac{\pi}{2p} \quad (15.56)$$

Noting that

$$f(t) = \int_0^\infty \frac{\sin(tx)}{x} dx = \mathcal{L}^{-1} \left[ \mathcal{L} \left[ f(t) \right] \right] = \mathcal{L}^{-1} \left[ \mathcal{L} \left[ \int_0^\infty \frac{\sin(tx)}{x} dx \right] \right]$$

(15.57)

using the the result we found the Laplace transform, we find

$$f(t) = \int_0^\infty \frac{\sin(tx)}{x} dx = \frac{\pi}{2p} \mathcal{L}^{-1} \left[ \frac{1}{p} \right] = \frac{\pi}{2} \quad \text{for } t > 0 \quad (15.58)$$

Noting that for

$$\sin(-tx) = -\sin(tx) \quad (15.59)$$

one can write

$$f(t) = \int_0^\infty \frac{\sin(tx)}{x} dx = -\frac{\pi}{2} \quad \text{for } t < 0. \quad (15.60)$$
For $t = 0$, obviously, we find
\[ f (t) = \int_{0}^{\infty} \frac{\sin (tx)}{x} \, dx = 0. \] (15.61)

Therefore we can write
\[ \int_{0}^{\infty} \frac{\sin (tx)}{x} \, dx = \begin{cases} \frac{\pi}{2} & t > 0 \\ 0 & t = 0 \\ -\frac{\pi}{2} & t < 0 \end{cases} \] (15.62)

### 15.2 Applications of Laplace transforms

In this section we will see applications of Laplace Transforms to solving differential equations. To this end, we recall the following "Table of Laplace Transforms" we have already determined for the basic functions

<table>
<thead>
<tr>
<th>$f (t)$</th>
<th>$1$</th>
<th>$e^{\omega t}$</th>
<th>$\sin (\omega t)$</th>
<th>$\cos (\omega t)$</th>
<th>$\sinh (\omega t)$</th>
<th>$\cosh (\omega t)$</th>
<th>$U (t - a)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L [f (t)]$</td>
<td>$\frac{1}{p}$</td>
<td>$\frac{1}{p - \omega}$</td>
<td>$\frac{\omega}{p^2 + \omega^2}$</td>
<td>$\frac{\omega}{p^2 + \omega^2}$</td>
<td>$\frac{p}{p^2 - \omega^2}$</td>
<td>$\frac{p}{p^2 - \omega^2}$</td>
<td>$e^{-\omega p}$</td>
</tr>
</tbody>
</table>

and the Laplace Transform for the $n^{th}$ order derivative of a continuous function, $f (t)$

\[
L [f^{(n)} (t)] = L \left[ \frac{d^n f (t)}{dt^n} \right] = p^n L [f (t)] - p^{n-1} f (0) - p^{n-2} \frac{df (0)}{dt} - \cdots - \frac{d^{n-1} f (0)}{dt^{n-1}}. \] (15.63)

**Example 14.5** Solve the following differential equation given the initial conditions, $y (0) = 1$ and $y' (0) = 0$

\[
\frac{d^2 y}{dx^2} + y (x) = 1. \] (15.64)

**Solution:** The Laplace transform of the differential equation can be written as

\[
L \left[ \frac{d^2 y}{dx^2} \right] + L [y (x)] = L [1] \] (15.65)

and applying the results we determined earlier, we have

\[
L [1] = \frac{1}{p}, L \left[ \frac{d^2 y}{dx^2} \right] = p^2 L [y (x)] - p y (0) - y' (0) \] (15.66)

so that using the given initial conditions

\[
y (0) = 1, y' (0) = 0 \] (15.67)

we obtain

\[
L \left[ \frac{d^2 y}{dx^2} \right] = p^2 L [y (x)] - p. \] (15.68)
Then the differential equation becomes
\[ p^2 \mathcal{L} [y(x)] - p + \mathcal{L} [y(x)] = \frac{1}{p} \Rightarrow (p^2 + 1) \mathcal{L} [y(x)] = \frac{1}{p} + p \]
\[ \Rightarrow \mathcal{L} [y(x)] = \frac{1}{p} \quad (15.69) \]

There follows that
\[ y(x) = \mathcal{L}^{-1} \left[ \frac{1}{p} \right] = 1 \quad (15.70) \]

Example 14.6 Solve the following differential equation given the initial conditions, \( y(0) = 2 \) and \( y'(0) = -1 \)
\[ \frac{d^2y}{dt^2} + 3 \frac{dy}{dt} + 2y(t) = 2e^{-t}. \quad (15.71) \]

**Solution:** Take the Laplace transform of each term and apply the given initial conditions
\[ \mathcal{L} \left[ \frac{d^2y}{dt^2} \right] = p^2 \mathcal{L} [y(t)] - py(0) - \frac{dy(0)}{dt} = p^2 \mathcal{L} [y(t)] - 2p + 1, \]
\[ \mathcal{L} \left[ \frac{dy}{dt} \right] = p\mathcal{L} [y(t)] - y(0) = p\mathcal{L} [y(t)] - 2, \]
\[ \mathcal{L} [e^{-t}] = \mathcal{L} [e^{-t}] = \frac{1}{p+1}. \quad (15.72) \]

The Laplace transform of the differential equation
\[ \mathcal{L} \left[ \frac{d^2y}{dt^2} \right] + 3\mathcal{L} \left[ \frac{dy}{dt} \right] + 2\mathcal{L} [y(t)] = 2\mathcal{L} [e^{-t}] \quad (15.73) \]
can then be written as
\[ p^2 \mathcal{L} [y(t)] - 2p + 1 + 3(p\mathcal{L} [y(t)] - 2) + 2\mathcal{L} [y(t)] = 2 \frac{1}{p+1} \]
\[ \Rightarrow (p^2 + 3p + 2) \mathcal{L} [y(t)] - 2p - 5 = \frac{2}{p+1} \]
\[ \Rightarrow (p^2 + 3p + 2) \mathcal{L} [y(t)] = 5 + 2p + \frac{2}{p+1} \quad (15.74) \]
Upon simplifying this expression, we find

\[ \mathcal{L}[y(t)] = \frac{5 + 2p}{(p + 2)(p + 1)} + \frac{2}{(p + 2)(p + 1)^2} \]

\[ \Rightarrow \mathcal{L}[y(t)] = \frac{2p^2 + 7p + 7}{(p + 2)(p + 1)^2}. \] (15.75)

Noting that

\[ \frac{2p^2 + 7p + 7}{(p + 2)(p + 1)^2} = \frac{A}{p + 2} + \frac{B}{p + 1} + \frac{C}{(p + 1)^2} \]

\[ = \frac{A(p^2 + 2p + 1) + B(p^2 + 3p + 2) + C(p + 2)}{(p + 2)(p + 1)^2} \]

\[ = \frac{(A + B)p^2 + (2A + 3B + C)p + A + 2B + 2C}{(p + 2)(p + 1)^2}. \] (15.76)

There follows that

\[ A + B = 2, 2A + 3B + C = 7, A + 2B + 2C = 7 \] (15.77)

**Mathematica result:**

**In:**

\[ eqns = \{ A + B == 2, A + 3B + C == 7, A + 2B + 2C == 7 \} \]

**In:**

\[ \text{Solve} \{eqns, \{A, B, C\}\} \]

**Out:**

\[ \{\{A \rightarrow 1, B \rightarrow 1, C \rightarrow 2\}\} \]

we find

\[ A = 1, B = 1, C = 2. \]

Then the Laplace transform for \( y(t) \) can be expressed as

\[ \mathcal{L}[y(t)] = \frac{2p^2 + 7p + 7}{(p + 2)(p + 1)^2} = \frac{1}{p + 2} + \frac{1}{p + 1} + \frac{2}{(p + 1)^2}. \] (15.78)
There follows that
\[ y(t) = \mathcal{L}^{-1} [\mathcal{L} [y(t)]] \]
\[ = \mathcal{L}^{-1} \left[ \frac{1}{p + 2} \right] + \mathcal{L}^{-1} \left[ \frac{1}{p + 1} \right] + \mathcal{L}^{-1} \left[ \frac{2}{(p + 1)^2} \right] \]  
(15.79)

Noting that
\[ \mathcal{L}^{-1} \left[ \frac{1}{p + 2} \right] = e^{-2t}, \mathcal{L}^{-1} \left[ \frac{1}{p + 1} \right] = e^{-t} \]  
(15.80)

and
\[ \frac{2}{(p + 1)^2} = -2 \frac{d}{d\lambda} \left( \frac{1}{p + \lambda} \right) \bigg|_{\lambda=1} \]
\[ \Rightarrow \mathcal{L}^{-1} \left[ \frac{2}{(p + 1)^2} \right] = \mathcal{L}^{-1} \left[ -2 \frac{d}{d\lambda} \left( \frac{1}{p + \lambda} \right) \bigg|_{\lambda=1} \right] \]
\[ = -2 \frac{d}{d\lambda} \mathcal{L}^{-1} \left[ \frac{1}{p + \lambda} \right] \bigg|_{\lambda=1} = -2 \frac{de^{-\lambda t}}{d\lambda} \bigg|_{\lambda=1} = -2t e^{-t} \]
\[ \Rightarrow \mathcal{L}^{-1} \left[ \frac{2}{(p + 1)^2} \right] = 2te^{-t} \]  
(15.81)

the solution to the differential equation which is given by
\[ y(t) = \mathcal{L}^{-1} \left[ \frac{1}{p + 2} \right] + \mathcal{L}^{-1} \left[ \frac{1}{p + 1} \right] + \mathcal{L}^{-1} \left[ \frac{2}{(p + 1)^2} \right] \]  
(15.82)

becomes
\[ y(t) = e^{-2t} + e^{-t} + 2te^{-t} = e^{-2t} [1 + e^t (1 + 2t)] \]  
(15.83)

**Mathematica result:**

In:

\[
\text{InverseLaplaceTransform}\left[\frac{5 + 2 p + \frac{2}{1+p}}{2 + 3 p + p^2}, p, t\right]
\]

Out:

\[ e^{-2 \cdot t} (1 + e^t (1 + 2t)) \]

By directly solving the differential equation using mathematica:

**Mathematica result:**

In:

\[
\text{Simplify}[\text{DSolve}\left\{y''[t] + 3y'[t] + 2 y[t] == 2 \text{Exp}[-t], y[t] == 2, y'[0] == -1\right\}, y, t]
\]
Out:
\[
\{ y \rightarrow \text{Function} \left[ \{ t \}, e^{-2 \cdot t} \left( 1 + e^{t} + 2 \cdot e^{2 \cdot t} \right) \right] \}
\]

**Example 14.7** Consider the circuit shown in the figure below. At \( t = 0 \) the switch \( S \) is closed to position \( A \). Find the resulting current, \( I(t) \). Use \( R = 10\Omega, L = 2H, \varepsilon (t) = 50 \sin (5t) \)

Applying Kirchoff’s voltage rule, we can write
\[
\varepsilon (t) - IR - L \frac{dI}{dt} = 0 \Rightarrow 50 \sin (5t) - 10I - 2 \frac{dI}{dt} = 0. \tag{15.84}
\]

The Laplace transform can be written as
\[
25\mathcal{L} [\sin (5t)] - 5[I] - \mathcal{L} \left[ \frac{dI}{dt} \right] = 0. \tag{15.85}
\]

Using
\[
\mathcal{L} [\sin (5t)] = \frac{5}{p^2 + 25}, \quad \mathcal{L} \left[ \frac{dI}{dt} \right] = p\mathcal{L} [I(t)] - I(0) \tag{15.86}
\]
we find
\[
\Rightarrow \mathcal{L} [I(t)] = \frac{125}{p^2 + 25} - 5\mathcal{L} [I] - p\mathcal{L} [I(t)] + I(0) = 0
\]
\[
\Rightarrow \mathcal{L} [I(t)] = \frac{125}{p^2 + 25} + I(0) \left( \frac{125}{(p^2 + 25)(p + 5)} + \frac{I(0)}{p + 5} \right) \tag{15.87}
\]
Noting that
\[
\frac{125}{(p^2 + 5^2)(p + 5)} = \frac{A}{p + 5} + \frac{BP + C}{p^2 + 5^2} = \frac{Ap^2 + 25A + BP^2 + 5Bp + Cp + 5C}{p^2 + 5^2}
\]
\[
= \frac{(A + B)p^2 + (5B + C)p + 25A + 5C}{p^2 + 5^2}
\]
\[A + B = 0, 5B + C = 0, 25A + 5C = 125\]
\[\Rightarrow A = \frac{5}{2}, B = -\frac{5}{2}, C = \frac{25}{2}\]
(15.88)

one can express
\[
\frac{125}{(p^2 + 5^2)(p + 5)} = \frac{5}{2} \frac{1}{p + 5} - \frac{5}{2} \frac{p}{p^2 + 5^2} + \frac{25}{2} \frac{1}{p^2 + 5^2}
\]
(15.89)

Just to check using Mathematica:

\[\text{In}[23]= \text{eqns} = \{A + B == 0, 5 B + C == 0, 25 A + 5 C == 125\};\]

\[\text{In}[24]= \text{Solve[eqns, \{A, B, C\}]}\]

\[\text{Out}[24]= \{\{A -> \frac{5}{2}, B -> -\frac{5}{2}, C -> \frac{25}{2}\}\}\]

Using these results, we can write
\[
\mathcal{L}[I(t)] = \left(\frac{5}{2} + I(0)\right) \frac{1}{p + 5} - \frac{5}{2} \frac{p}{p^2 + 5^2} + \frac{25}{2} \frac{1}{p^2 + 5^2}
\]
(15.90)

so that the inverse Laplace transform gives
\[
I(t) = \mathcal{L}^{-1}[\mathcal{L}[I(t)]] = \left(\frac{5}{2} + I(0)\right) \mathcal{L}^{-1}\left[\frac{1}{p + 5}\right] - \frac{5}{2} \mathcal{L}^{-1}\left[\frac{p}{p^2 + 5^2}\right] + \frac{5}{2} \mathcal{L}^{-1}\left[\frac{5}{p^2 + 5^2}\right]
\]
(15.91)

Referring to our "Table of Laplace's Transform"

<table>
<thead>
<tr>
<th>(f(t))</th>
<th>1</th>
<th>(e^{\omega t})</th>
<th>(\sin(\omega t))</th>
<th>(\cos(\omega t))</th>
<th>(\sinh(\omega t))</th>
<th>(\cosh(\omega t))</th>
<th>(U(t-a))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathcal{L}[f(t)])</td>
<td>(\frac{1}{p})</td>
<td>(\frac{1}{p - \omega})</td>
<td>(\frac{\omega}{p^2 + \omega^2})</td>
<td>(\frac{p}{p^2 - \omega^2})</td>
<td>(\frac{\omega}{p^2 - \omega^2})</td>
<td>(\frac{p}{p^2 - \omega^2})</td>
<td>(\frac{e^{\omega t}}{p})</td>
</tr>
</tbody>
</table>

we find
\[
I(t) = \left(\frac{5}{2} + I(0)\right) e^{-5t} - \frac{5}{2} \cos(5t) + \frac{5}{2} \sin(5t).
\]
(15.92)

In RL circuit the inductor acts as an open circuit at the initial time and we can set, \(I(0) = 0\), and the current becomes
\[
I(t) = \frac{5}{2} e^{-5t} + \frac{5}{2} |\sin(5t) - \cos(5t)|
\]
(15.93)

Mathematica result:
15.2. APPLICATIONS OF LAPLACE TRANSFORMS

In:

\[ DSolve \left[ 50 \sin(5t) - 10 I'[t] - 2I'[t] == 0, I, t \right], \]

Out:

\[ \left\{ I \rightarrow Function \left[ \left\{ t \right\}, e^{-5t}C[1] + \frac{5}{2}(-\cos[5t] + \sin[5t]) \right] \right\} \]

Note that we can write

\[ \left( \frac{5}{2} + I(0) \right) e^{-5t} = \left( \frac{5}{2} + I(0) \right) e^{-5t} = c[1]e^{-5t} \quad (15.94) \]

Example 14.8 Damped Harmonic oscillator: Consider a mass \( m \) oscillating under the influence of a spring, spring constant \( k \) and damped by a friction force which is proportional to the velocity (\( F_f = -bv \)). Assuming that the particle starts from rest at \( x(0) = x_0, \ x'(0) = 0 \). Find the equation of motion for the mass \( m \) and determine the position of the mass as function of time using Laplace transformation.

Solution:

\[ x(t) = x_0 e^{(-b/2m)t} \left( \cos(\omega_1 t) + \frac{b}{2m\omega_1} \sin(\omega_1 t) \right) \quad (15.95) \]

where

\[ \omega_1^2 = \frac{k}{m} - \frac{b^2}{4m^2} \quad (15.96) \]
Chapter 16

Fourier series and transform

In this chapter we will cover Fourier Series and Transform.

16.1 Introduction to Fourier Series

Periodic Functions: A function $f(x)$ is said to be a periodic function when the function repeats itself after a period $T$, which means

$$f(x) = f(x + T)$$

for all $x$.

Average Values: The average value of the function $f(x)$ in the interval, $a < x < b$, is given by

$$\langle f(x) \rangle = \frac{1}{b-a} \int_a^b f(x) \, dx$$

for a periodic function with a period $T$, we can write the average value in one period as

$$\langle f(x) \rangle = \frac{1}{T} \int_{-T/2}^{T/2} f(x) \, dx.$$

Periodic Motion: Periodic motion is when the motion of an object continually repeats itself. This can be repeatedly moving back and forth or it could be moving in a circular orbit or rotation. Since the Law of Inertia states that an object moves in a straight line unless acted upon by a force, periodic motion requires some sort of force to create this special type of motion. Characteristics of periodic motion are the velocity of the object, the period of motion and the amplitude of the motion. Example of Periodic motion include circular motion (e.g. when you swing an object around you that is held on a rope or string, motion of the planets around the sun) and back-and-forth motion (e.g. a bouncing ball, pendulum and spring).
Example 16.1 Compute \( \langle f(x) \rangle \) for \( f(x) = \sin(x) \) over the interval \( 0 < x < \pi \).

Solution: Using the expression

\[
\langle f(x) \rangle = \frac{1}{b-a} \int_a^b f(x) \, dx
\]  

(16.1)

we find

\[
\langle f(x) \rangle = \frac{1}{\pi} \int_0^\pi \sin(x) \, dx = -\frac{1}{\pi} \cos(x) \bigg|_0^\pi \Rightarrow \langle f(x) \rangle = \frac{2}{\pi}.
\]  

(16.2)

Example 16.2 Compute \( \langle f(x) \rangle \) for \( f(x) = \sin(x) \) over the interval \( -\pi < x < \pi \). That is, find \( \langle f(x) \rangle \) over one complete period.

Solution: Here we have changed the interval to \( -\pi < x < \pi \) and the average should be expressed as

\[
\langle f(x) \rangle = \frac{1}{\pi - (-\pi)} \int_{-\pi}^{\pi} \sin(x) \, dx = -\frac{1}{2\pi} \cos(x) \bigg|_{-\pi}^{\pi} = 0
\]  

(16.3)

Example 16.3 The current in an RLC circuit with an oscillating voltage source ("emf") of the form

\[
V_S(t) = V_m \sin(\omega t)
\]  

(16.4)

is given by

\[
I(t) = I_m \sin(\omega t - \phi),
\]  

(16.5)

where \( V_m = I_m Z \). The quantity \( Z \), called the total impedance of the circuit, is given by

\[
Z = \sqrt{R^2 + \left( \omega L - \frac{1}{\omega C} \right)^2}
\]  

(16.6)

for a series circuit. Find an expression for the root-mean-square current,

\[ I_{rms}, \]  

over one full period.
16.1. INTRODUCTION TO FOURIER SERIES

Solution: The root mean square current, \( I_{rms} \), is the square root of the average value for \( I^2(t) \) in one full period, \( T \). That means

\[
I_{rms} = \sqrt{\langle I^2(t) \rangle}.
\]  

So first we need to find \( \langle I^2(t) \rangle \). To this end, we recall the angular frequency \( \omega = 2\pi f = \frac{2\pi}{T} \), and the period, \( T \), can be expressed in terms of the angular frequency, \( \omega \), as

\[
T = \frac{2\pi}{\omega} \quad \Rightarrow \quad \frac{1}{T} = \frac{\omega}{2\pi}.
\]  

Then the average, \( \langle I^2(t) \rangle \) in the interval, \( 0 < t < T \), given by

\[
\langle I^2(t) \rangle = \frac{1}{T} \int_0^T I^2(t) \, dt
\]  

can be expressed as

\[
\langle I^2(t) \rangle = \frac{\omega}{2\pi} \int_0^T [I_m \sin(\omega t - \varphi)]^2 \, dt.
\]  

Introducing the variable

\[
x = \omega t - \varphi \Rightarrow \frac{dx}{\omega} = dt
\]  

and

\[
t = 0 \Rightarrow x = -\varphi,
\]

\[
t = T \Rightarrow x = \omega T - \varphi = \frac{2\pi}{\omega} - \varphi = 2\pi - \varphi
\]  

we have

\[
\langle I^2(t) \rangle = \frac{\omega}{2\pi} \int_{-\varphi}^{2\pi-\varphi} I_m^2 \sin^2(x) \left( \frac{dx}{\omega} \right) = \frac{I_m^2}{2\pi} \int_{-\varphi}^{2\pi-\varphi} \sin^2(x) \, dx
\]  

Noting that

\[
\int \sin^2(x) \, dx = \frac{x}{2} - \frac{1}{4} \sin(2x) = \frac{x}{2} - \frac{1}{2} \sin(x) \cos(x)
\]  

one can write

\[
\langle I^2(t) \rangle = \left[ \frac{x}{2} - \frac{1}{2} \sin(x) \cos(x) \right]_{-\varphi}^{2\pi-\varphi}
\]  

which gives

\[
\langle I^2(t) \rangle = \left[ \frac{I_m^2}{2\pi} \frac{2\pi - \varphi}{2} - \sin(2\pi - \varphi) \cos(2\pi - \varphi) \right]
\]

\[
\quad - \frac{I_m^2}{2\pi} \left[ \frac{-\varphi}{2} - \sin(-\varphi) \cos(-\varphi) \right].
\]  

(16.15)
CHAPTER 16. FOURIER SERIES AND TRANSFORM

Noting that
\[
\sin(-\varphi) = -\sin(\varphi), \quad \cos(-\varphi) = \cos(\varphi)
\]
\[
\sin(\varphi \pm 2\pi) = \sin(\varphi), \quad \cos(\varphi \pm 2\pi) = \cos(\varphi)
\]
we find
\[
\langle I^2(t) \rangle = \frac{I_m^2}{2}.
\]  
(16.17)

Then the root mean square current, \( I_{rms} \), is the square root of the average value for \( I^2(t) \) in one full period, \( T \), becomes
\[
I_{rms} = \sqrt{\langle I^2(t) \rangle} = \frac{I_m}{\sqrt{2}}.
\]
(16.18)

16.2 The Fourier Series Expansion of a Periodic Function

Any function of periodicity \( 2\pi \) can be expanded in a Fourier series of the form
\[
f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left[ a_n \cos(nx) + b_n \sin(nx) \right]
\]
(16.19)

where the coefficients are given by
\[
a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, dx, \quad n = 0, 1, 2, 3, \ldots
\]
(16.20)
\[
b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx, \quad n = 1, 2, 3, \ldots
\]
(16.21)

Using Euler’s formula
\[
\cos(nx) = \frac{e^{inx} + e^{-inx}}{2}, \quad \sin(nx) = \frac{e^{inx} - e^{-inx}}{2i}
\]
(16.22)

we may
\[
f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left[ a_n \left( \frac{e^{inx} + e^{-inx}}{2} \right) + b_n \left( \frac{e^{inx} - e^{-inx}}{2i} \right) \right]
\]
\[
= \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left[ \left( \frac{a_n - ib_n}{2} \right) e^{inx} + \left( \frac{a_n + ib_n}{2} \right) e^{-inx} \right]
\]

Replacing \( n \) by \( -n \) in the second series and noting that
\[
a_0 = a_0 e^{i\theta(x)}
\]
(16.23)
one can write
\[
f(x) = \frac{1}{2} a_0 e^{i0\pi x} + \sum_{n=1}^{\infty} \left( a_n - ib_n \right) e^{inx} + \sum_{n=-1}^{\infty} \left( a_n + ib_n \right) e^{inx} = \sum_{n=-\infty}^{\infty} c_n e^{inx},
\]
(16.24)

where
\[
c_n = \begin{cases} \\
\frac{1}{2}a_0 & n = 0 \\
\frac{a_n - ib_n}{2} & n \geq 1 \\
\frac{a_n + ib_n}{2} & n \leq -1
\end{cases}
\]
(16.25)

More concisely the expression for \(c_n\) can be determined from
\[
\begin{align*}
a_n - ib_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \left[ \cos(nx) - i \sin(nx) \right] dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \\
n &= 0, 1, 2, 3, \ldots, \\
a_n + ib_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \left[ \cos(nx) + i \sin(nx) \right] dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{inx} dx,
n &= 1, 2, 3, \ldots
\end{align*}
\]
(16.26)

which leads to
\[
c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad \text{where } n = \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots
\]

and the Fourier series expansion for the periodic function \(f(x)\) is given by
\[
f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx},
\]
(16.27)

Example 16.4 Find the Fourier series expansion for the function
\[
f(x) = \begin{cases} \\
1 & \text{for } 0 < x < \pi \\
-1 & \text{for } -\pi < x < 0
\end{cases}
\]
(16.28)

which is repeated indefinitely with a period of \(2\pi\).

Solution: Using
\[
a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx, \quad n = 0, 1, 2, 3, \ldots
\]
(16.29)
we have

\[ a_n = \frac{1}{\pi} \left[ -\int_{0}^{\pi} \cos (nx) \, dx + \int_{0}^{\pi} \cos (nx) \, dx \right] \]

\[ = \frac{1}{\pi} \left[ \int_{0}^{\pi} \cos (-nx) \, d(-x) + \int_{0}^{\pi} \cos (nx) \, dx \right] \\
= \frac{1}{\pi} \left[ -\int_{0}^{\pi} \cos (nx) \, dx + \int_{0}^{\pi} \cos (nx) \, dx \right] \\
\Rightarrow a_n = 0, \quad n = 0, 1, 2, 3... \quad (16.30) \]

And using

\[ b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin (nx) \, dx, \quad n = 1, 2, 3... \]

\[ = \frac{1}{\pi} \left[ -\int_{-\pi}^{0} \sin (nx) \, dx + \int_{0}^{\pi} \sin (nx) \, dx \right] \]

\[ = \frac{1}{\pi} \left[ -\int_{-\pi}^{0} \sin (-nx) \, d(-x) + \int_{0}^{\pi} \sin (nx) \, dx \right] \\
= \frac{1}{\pi} \left[ \int_{0}^{\pi} \sin (nx) \, dx + \int_{0}^{\pi} \sin (nx) \, dx \right] \\
= \frac{2}{\pi} \int_{0}^{\pi} \sin (nx) \, dx = \frac{2 \cos (nx)}{\pi n} \bigg|_{0}^{\pi} = \frac{2(1 - (-1)^n)}{\pi n} \\
\Rightarrow a_n = \left\{ \begin{array}{ll}
\frac{4}{\pi n} & n = 1, 3, 5... \\
\frac{4}{\pi (2n + 1)} & n = 2, 4, 6... \end{array} \right\}, \quad n = 0, 1, 2, 3... \quad (16.31) \]

Therefore, the Fourier series expansion of

\[ f(x) = \left\{ \begin{array}{ll}
1 & \text{for } 0 < x < \pi \\
-1 & \text{for } -\pi < x < 0
\end{array} \right\} \]

is given by

\[ f(x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin [(2n + 1)x]}{2n + 1}. \quad (16.33) \]
16.3 Dirichlet Conditions

The Dirichlet conditions are sufficient conditions for a real-valued, periodic function \( f(x) \) to be equal to the sum of its Fourier series at each point where \( f \) is continuous. Moreover, the behavior of the Fourier series at points of discontinuity is determined as well. These conditions are named after Johann Peter Gustav Lejeune Dirichlet. The conditions are:

(a) \( f(x) \) must have a finite number of extrema (maxima, minima) in any given interval (period)

(b) \( f(x) \) must have a finite number of discontinuities in any given interval

(c) \( f(x) \) must be absolutely integrable over a period.

(d) \( f(x) \) must be bounded within that period (that is, it doesn’t become infinite)

If a function has a periodicity of \( 2\pi \), has a finite number of maxima, minima, and discontinuities within a single period, and if it is bounded within that period (that is, it doesn’t become infinite), then the Fourier series will converge
to the given function for all values of \( x \) where the function is continuous, and will converge to the average value of the function at any value of \( x \) where the function has a discontinuity.

**Dirichlet’s Theorem for 1-Dimensional Fourier Series:** If \( f(x) \) satisfies Dirichlet conditions, the Fourier series will converge to the given function for all values of \( x \) where the function is continuous, and will converge to the average value of the function at any value of \( x \) where the function has a discontinuity. This means

\[
\frac{1}{2}a_0 + \sum_{n=1}^{\infty} [a_n \cos (nx) + b_n \sin (nx)] = \sum_{n=-\infty}^{\infty} c_n e^{inx} = \frac{1}{2} (f(x_+) + f(x_-))
\]

(16.34)

where the notation

\[
f(x_+) = \lim_{y \to x_+} f(y) \quad f(x_-) = \lim_{y \to x_-} f(y)
\]

(16.35)

denotes the right/left limits of \( f \).

**Example 15.5** Verify Dirichlet’s Theorem for the function in the previous example

\[
f(x) = \begin{cases} 
1 & \text{for } 0 < x < \pi \\
-1 & \text{for } -\pi < x < 0
\end{cases}
\]

(16.36)

**Solution:** We note that

(a) \( f(x) \) have a finite number of extrema (maxima, minima) in any given interval, \(-\pi < x < \pi\).

(b) \( f(x) \) has one discontinuity at \( x = 0 \) in the interval \(-\pi < x < \pi\).

(c) \( f(x) \) is absolutely integrable over \(-\pi < x < \pi\).

(d) \( f(x) \) is bounded within \(-\pi < x < \pi\).

We must be able to show that

\[
\sum_{n=-\infty}^{\infty} c_n e^{inx} = \frac{1}{2} (f(x_+) + f(x_-)).
\]

(16.37)

Recalling that

\[
\sum_{n=-\infty}^{\infty} c_n e^{inx} = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\sin [(2n+1)x]}{2n+1}
\]

(16.38)

for \( x = 0 \), we have

\[
\sum_{n=-\infty}^{\infty} c_n e^{inx} = 0.
\]

(16.39)
Noting that
\[ f(x+) = 1, \quad f(x-) = -1 \] (16.40)
we find
\[ \frac{1}{2} (f(x+) + f(x-)) = 0. \] (16.41)
Therefore, we can conclude that
\[ \sum_{n=-\infty}^{\infty} c_n e^{inx} = \frac{1}{2} (f(x+) + f(x-)). \] (16.42)

\subsection{16.4 Fourier series with spatial and temporal arguments}

For \( x \) unit less (or in radians) and periodic on \( 2\pi \), \( f(x + 2\pi) = f(x) \),
\[ f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left[ a_n \cos(nx) + b_n \sin(nx) \right] \] (16.43)
where the coefficients are determined using
\[ a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, dx, \quad n = 0, 1, 2, 3... \]
\[ b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx, \quad n = 1, 2, 3... \] (16.44)

\textit{Spatial Arguments:} If \( x \) represents a position in space, and \( f(x + \lambda) = f(x) \),
then, for \( \lambda = 2\pi/\lambda \),
\[ f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left[ a_n \cos(n\lambda x) + b_n \sin(n\lambda x) \right] \] (16.45)
where the coefficients
\[ a_n = \frac{2}{\lambda} \int_{-\lambda/2}^{\lambda/2} f(x) \cos(n\lambda x) \, dx, \quad n = 0, 1, 2, 3... \]
\[ b_n = \frac{2}{\lambda} \int_{-\lambda/2}^{\lambda/2} f(x) \sin(n\lambda x) \, dx, \quad n = 1, 2, 3... \] (16.46)

\textit{Temporal Arguments:} If \( t \) represents time, and \( f(t + T) = f(t) \), then, for \( \omega = 2\pi/T \),
\[ f(t) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left[ a_n \cos(n\omega t) + b_n \sin(n\omega t) \right] \] (16.47)
where the coefficients

\[ a_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos(n \omega t) \, dt, \quad n = 0, 1, 2, 3, \ldots \]

\[ b_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin(n \omega t) \, dt, \quad n = 1, 2, 3, \ldots \] (16.48)

Example 15.6 A voltage signal in the form of a half-rectified sine-wave is fed into a linear circuit whose response to a simple sinusoidal signal is known. In order to determine the response of the circuit to the input signal, we wish to find the Fourier series expansion of that signal. Assume that the input signal is given by

\[ V(t) = \begin{cases} 0 & -T/2 < t < 0 \\ \Delta V_m \sin(\omega t) & 0 < t < T/2 \end{cases} \] (16.49)

**Hint:** You may find the following integrals useful.

\[ \int_{-T/2}^{T/2} \cos(m \omega t) \sin(\omega t) \, dt = \frac{\cos[(1-m) \omega t]}{2(1-m) \omega} - \frac{\cos[(1+m) \omega t]}{2(1+m) \omega} \] (16.50)

**Solution:** Using Fourier expansion coefficients for temporal arguments

\[ a_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos(n \omega t) \, dt, \quad n = 0, 1, 2, 3, \ldots \] (16.51)

\[ b_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin(n \omega t) \, dt, \quad n = 1, 2, 3, \ldots \] (16.52)

for the given function one can write

\[ a_n = \frac{2 \Delta V_m}{T} \int_{0}^{T/2} \cos(n \omega t) \sin(\omega t) \, dt, \quad b_n = \frac{2 \Delta V_m}{T} \int_{0}^{T/2} \sin(n \omega t) \sin(\omega t) \, dt. \] (16.53)

Applying the trigonometric relation

\[ \sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \cos \alpha \sin \beta \]

\[ \cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta \] (16.54)
we can write

\[
\cos (n \omega t) \sin (\omega t) = \frac{1}{2} \{ \sin [(n+1) \omega t] - \sin [(n-1) \omega t] \} \quad (16.55)
\]

\[
\sin (n \omega t) \sin (\omega t) = \frac{1}{2} \{ \cos [(n-1) \omega t] - \cos [(n+1) \omega t] \}
\]

so that

\[
a_n = \frac{\Delta V_m}{T} \left\{ \int_0^{T/2} \sin [(n+1) \omega t] dt - \int_0^{T/2} \sin [(n-1) \omega t] dt \right\}
\]

\[
b_n = \frac{\Delta V_m}{T} \left\{ \int_0^{T/2} \cos [(n-1) \omega t] dt - \int_0^{T/2} \cos [(n+1) \omega t] dt \right\} \quad (16.56)
\]

For \( n = 0 \)

\[
a_0 = \frac{2\Delta V_m}{T} \int_0^{T/2} \sin [\omega t] dt = \frac{2\Delta V_m}{T \omega} = \frac{\Delta V_m}{\pi} \quad (16.57)
\]

for \( n = 1 \)

\[
a_1 = \frac{\Delta V_m}{T} \int_0^{T/2} \sin [2 \omega t] dt = \frac{\Delta V_m}{2T \omega} = \frac{\Delta V_m}{\pi} \quad (16.58)
\]

\[
b_1 = \frac{\pi \Delta V_m}{2T} - \frac{\Delta V_m}{T} \int_0^{T/2} \cos [2 \omega t] dt = \frac{\pi \Delta V_m}{2T} - \frac{\Delta V_m}{2T \omega} = \frac{\Delta V_m}{4\pi} \quad (16.59)
\]

where we used

\[
\omega = \frac{2\pi}{T} \Rightarrow T \omega = 2\pi \quad (16.60)
\]

For \( n > 1 \)

\[
a_n = \frac{\Delta V_m}{T} \left\{ \frac{1}{n+1} \frac{\cos [(n+1) \omega t/2]}{n + 1} - \frac{1}{n-1} + \frac{\cos [(n-1) \omega T/2]}{n-1} \right\}^{T/2}
\]

\[
= \frac{\Delta V_m}{2\pi} \left\{ -\frac{2}{n^2 - 1} + \frac{\cos [(n+1) \pi]}{n+1} - \frac{\cos [(n-1) \pi]}{n-1} \right\}
\]

\[
\Delta V_m \left\{ -\frac{2}{n^2 - 1} + \frac{\cos (n \pi)}{n+1} - \frac{\cos (n \pi)}{n-1} \right\}
\]

\[
= \frac{\Delta V_m}{\pi} \left\{ \frac{1}{n^2 - 1} + \frac{(-1)^n}{n^2 - 1} \right\}
\]

\[
\Rightarrow a_n = -\frac{\Delta V_m}{\pi} \frac{2}{(2n)^2 - 1}, \quad n = 1, 2, 3, \ldots \quad (16.61)
\]
and
\[ b_n = \frac{\Delta V_m}{T} \left\{ \int_0^{T/2} \cos [(n-1)\omega t] \, dt - \int_0^{T/2} \cos [(n+1)\omega t] \, dt \right\} \]
\[ = \frac{\Delta V_m}{\omega T} \left\{ \frac{\sin [(n-1)\omega t]}{n-1} - \frac{\sin [(n+1)\omega t]}{n+1} \right\} \bigg|_0^{T/2} \]
\[ \Rightarrow b_n = \frac{\Delta V_m}{2\pi} \left\{ \frac{\sin [(n-1)\pi]}{n-1} - \frac{\sin [(n+1)\pi]}{n+1} \right\} = 0 \quad (16.62) \]

Using the results
\[ a_0 = \frac{\Delta V_m}{\pi}, a_1 = \frac{\Delta V_m}{4\pi}, a_{2n} = -\frac{\Delta V_m}{\pi} \frac{2}{(2n)^2 - 1}, n = 1, 2, 3.. \]
\[ b_1 = \frac{\pi \Delta V_m}{2T} - \frac{\Delta V_m}{4\pi}, b_n = 0 \quad (16.63) \]
we can write the Fourier series as
\[ f(t) = \frac{\Delta V_m}{2\pi} + \frac{\Delta V_m}{4\pi} \cos(\omega t) + \left( \frac{\omega \Delta V_m}{4} - \frac{\Delta V_m}{4\pi} \right) \sin(\omega t) \]
\[ - \frac{2\Delta V_m}{\pi} \sum_{n=1}^{\infty} \frac{\cos(2n\omega t)}{4n^2 - 1}. \quad (16.64) \]

16.5 The Fourier Transform

The Fourier transform of a periodic or non periodic function, \( f(t) \), that satisfies all the Dirichlet conditions does exist and is given by
\[ \mathcal{F}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \, dt = F(\omega). \quad (16.65) \]

The Inverse Fourier Transform: We may develop the inverse Fourier Transform starting from the Fourier series expansion. Suppose we are considering the interval \([-T, T]\) instead of \([-\pi, \pi]\) the Fourier series expansion for the function \( f(t) \) can be expressed as
\[ f(t) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left[ a_n \cos\left(\frac{n\pi}{T} t\right) + b_n \sin\left(\frac{n\pi}{T} t\right) \right] \quad (16.66) \]
where the Fourier Coefficients are given by
\[ a_n = \frac{1}{T} \int_{-T}^{T} f(t') \cos\left(\frac{n\pi}{T} t'\right) \, dt', n = 0, 1, 2, 3... \]
\[ b_n = \frac{1}{T} \int_{-T}^{T} f(t') \sin\left(\frac{n\pi}{T} t'\right) \, dt', n = 1, 2, 3... \quad (16.67) \]
The resulting Fourier series can be written as

\[
f(t) = \frac{1}{T} \int_{-T}^{T} f(t') \, dt' + \frac{1}{T} \sum_{n=1}^{\infty} \cos \left( \frac{\pi n}{T} t \right) \int_{-T}^{T} f(t') \cos \left( \frac{\pi n}{T} t' \right) \, dt'
\]

\[
+ \frac{1}{T} \sum_{n=1}^{\infty} \sin \left( \frac{\pi n}{T} t \right) \int_{-T}^{T} f(t') \sin \left( \frac{\pi n}{T} t' \right) \, dt'
\]

\[
= \frac{1}{T} \int_{-T}^{T} f(t') \, dt' + \frac{1}{T} \sum_{n=1}^{\infty} \int_{-T}^{T} f(t') \left\{ \cos \left( \frac{\pi n}{T} t \right) \cos \left( \frac{\pi n}{T} t' \right) + \sin \left( \frac{\pi n}{T} t \right) \sin \left( \frac{\pi n}{T} t' \right) \right\} \, dt'
\]

\[
\Rightarrow f(t) = \frac{1}{T} \int_{-T}^{T} f(t') \, dt' + \sum_{n=1}^{\infty} \frac{1}{T} \int_{-T}^{T} f(t') \cos \left[ \frac{\pi n}{T} (t - t') \right] \, dt',
\]

where we used the double angle formula for cosine. Now introducing the variable defined by

\[
\omega_n = \frac{\pi n}{T} \Rightarrow \Delta \omega = \frac{\omega_{n+1} - \omega_n}{\pi} = \frac{1}{T}
\]

so that

\[
f(t) = \frac{1}{\pi} \Delta \omega \int_{-T}^{T} f(t') \, dt' + \frac{1}{\pi} \sum_{n=1}^{\infty} \Delta \omega \int_{-T}^{T} f(t') \cos (\omega_n (t - t')) \, dt'
\]

or

\[
f(t) = \frac{1}{\pi} \sum_{n=0}^{\infty} \Delta \omega \int_{-T}^{T} f(t') \cos (\omega_n (t - t')) \, dt'.
\]

We now let the parameter \( T \) approach infinity, transforming the interval \([-T, T]\) into the infinite interval \((-\infty, \infty)\). In this limit we have

\[
\Delta \omega \rightarrow d\omega, \omega_n \rightarrow \omega
\]

and if the function \( f \) satisfies the Dirichlet conditions, we can replace the summation by integration

\[
f(t) = \frac{1}{\pi} \int_{0}^{\infty} d\omega \int_{-\infty}^{\infty} f(t') \cos [\omega (t - t')] \, dt'.
\]

Noting that \( \cos [\omega (t - t')] \) is an even function and \( \sin [\omega (t - t')] \) is an odd function of \( \omega \), we have

\[
\int_{0}^{\infty} d\omega \int_{-\infty}^{\infty} f(t') \cos [\omega (t - t')] \, dt' = \frac{1}{2} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} f(t') \cos [\omega (t - t')] \, dt'
\]

\[
(16.74)
\]
and
\[ i \frac{1}{2} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} f(t') \sin[\omega(t - t')] dt' = 0 \] (16.75)

so that
\[ f(t) = \frac{1}{\pi} \int_{0}^{\infty} d\omega \int_{-\infty}^{\infty} f(t') \cos[\omega(t - t')] dt' \]
\[ = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} f(t') \{ \cos[\omega(t - t')] + i \sin[\omega(t - t')] \} dt' \]
\[ \Rightarrow f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} d\omega e^{i\omega t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t') e^{-i\omega t'} dt' \] (16.76)

We recall
\[ F(\omega) = \mathcal{F}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \] (16.77)

so that
\[ f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega = \mathcal{F}^{-1}[F(\omega)] \] (16.78)

which is the Inverse Fourier Transform.

**Note:** If \( f(t) \) has a discontinuity, then the inverse Fourier transform will return a value for \( f(t) \) that is at the midpoint of the discontinuity at the x-value at which the discontinuity occurs.

**Example 15.7** Find and describe the Fourier transform of the function \( f(t) \) defined by
\[ f(t) = \begin{cases} 
1 & |t| < 1 \\
0 & |t| > 1 
\end{cases} \] (16.79)

which is an even function of \( t \). This is the single slit diffraction problem of physical optics. The slit is described by \( f(t) \). The diffraction pattern amplitude is given by the Fourier transform \( F(\omega) \).

**Solution:** We recall
\[ \mathcal{F}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \] (16.80)

so that for the function given we can write
\[ \mathcal{F}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \left[ \frac{e^{-i\omega t}}{i\omega} \right]_{-1}^{1} \]
\[ = \frac{2}{\sqrt{2\pi}\omega} e^{i\omega} - e^{-i\omega} \Rightarrow \mathcal{F}[f(x)] = \frac{\sqrt{2}}{\pi} \frac{\sin(\omega)}{\omega} \] (16.81)
Note: The function

\[ f(t) = \frac{\sin(x)}{x} \]  

is known as the *sinc function* commonly expressed as \( \text{sinc}(x) \).

**Example 15.8** Find the Fourier Transform of a Gaussian function

\[ f(x) = e^{-ax^2} \]  

**Solution:** For the given function we can write

\[ \mathcal{F}[f(x)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(ax^2 + i\omega x)} \, dx \]  

and using

\[ a^2 x^2 + i\omega x = a^2 \left( x + \frac{i\omega}{2a^2} \right)^2 + \left( \frac{\omega}{2a^2} \right)^2 = a^2 \left( x + \frac{i\omega}{2a^2} \right)^2 + \frac{\omega^2}{4a^2} \]  

one can write

\[ \mathcal{F}[f(x)] = \frac{e^{-\frac{\omega^2}{4a^2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-a^2 \left( x + \frac{i\omega}{2a^2} \right)^2} \, dx. \]
To evaluate the integral
\[ I = \int_{-\infty}^{\infty} e^{-a^2(x + \frac{\omega}{2\pi})^2} \, dx \]  
\[(16.87)\]

consider the integral
\[ \int e^{-a^2z^2} \, dz \]  
\[(16.88)\]

for the rectangular closed loop with vertices \(-T, T, T + \frac{i\omega}{2\pi}, -T + \frac{i\omega}{2\pi}\). For this integral applying the residue theorem we can write
\[
\int e^{-a^2z^2} \, dz = \int_{-\infty}^{\infty} e^{-a^2x^2} \, dx + i \int_{-\infty}^{0} e^{-a^2(T+iy)^2} \, dy
\]
\[
+ \int_{0}^{\frac{\omega}{2\pi}} e^{-a^2(x + \frac{i\omega}{2\pi})^2} \, dx + i \int_{\frac{\omega}{2\pi}}^{T} e^{-a^2(-T+iy)^2} \, dy
\]
\[
= \int_{-T}^{T} e^{-a^2x^2} \, dx + i \int_{0}^{\frac{\omega}{2\pi}} e^{-a^2(T+iy)^2} \, dy + \int_{T}^{0} e^{-a^2(x + \frac{i\omega}{2\pi})^2} \, dx + i \int_{-\frac{\omega}{2\pi}}^{0} e^{-a^2(T-iy)^2} \, dy
\]
\[(16.89)\]

Noting that in the limit \(T \to \infty\),
\[
\int_{0}^{\frac{\omega}{2\pi}} e^{-a^2(T+iy)^2} \, dy = \int_{-\frac{\omega}{2\pi}}^{0} e^{-a^2(T-iy)^2} \, dy = 0
\]
\[(16.90)\]

and there is no poles inside the curve, we may write
\[
\int e^{-a^2z^2} \, dz \quad 0 = \int_{-\infty}^{\infty} e^{-a^2x^2} \, dx + \int_{-\infty}^{\infty} e^{-a^2(x + \frac{i\omega}{2\pi})^2} \, dx = \int_{-\infty}^{\infty} e^{-a^2x^2} \, dx
\]
\[
-\int_{-\infty}^{\infty} e^{-a^2(x + \frac{i\omega}{2\pi})^2} \, dx = 0 \Rightarrow \int_{-\infty}^{\infty} e^{-a^2(x + \frac{i\omega}{2\pi})^2} \, dx = \int_{-\infty}^{\infty} e^{-a^2x^2} \, dx
\]
\[(16.91)\]

We may write
\[
\int_{-\infty}^{\infty} e^{-a^2x^2} \, dx = \frac{1}{a} \int_{-\infty}^{\infty} e^{-y^2} \, dy
\]
\[(16.92)\]
so that applying the result

\[
\left( \int_{-\infty}^{\infty} e^{-y^2} \, dy \right)^2 = \int_{-\infty}^{\infty} e^{-u^2} \, du \int_{-\infty}^{\infty} e^{-v^2} \, dv = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(u^2+v^2)} \, du \, dv = \int_0^{2\pi} \int_0^{\infty} e^{-r^2} \, r \, dr \, d\theta
\]

\[
= 2\pi \int_0^{\infty} e^{-r^2} \, r \, dr = -\pi e^{-r^2} \bigg|_0^\infty = \pi \Rightarrow \int_{-\infty}^{\infty} e^{-y^2} \, dy = \sqrt{\pi}
\] (16.93)

In[1]:= \[\text{Exp}\left[-a^2 \cdot x^2\right] \, dx\]

Out[1]= If[\[\text{Re}[a^2] > 0, \frac{\sqrt{\pi}}{\sqrt{a^2}}, \text{Integrate}\left[e^{-a^2 \cdot x^2}, \{x, -\infty, \infty\}, \text{Assumptions} \to \text{Re}[a^2] \leq 0\right]\] we find

\[
\int_{-\infty}^{\infty} e^{-a^2 (x+\frac{\omega}{a^2})^2} \, dx = \frac{\sqrt{\pi}}{a}.
\] (16.94)

Therefore the Fourier integral transform

\[
\mathcal{F}[f(x)] = \frac{e^{-\frac{\omega^2}{4a^2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-a^2 x^2} \, dx
\] (16.95)

becomes

\[
\mathcal{F}[f(x)] = \frac{e^{-\frac{\omega^2}{4a^2}}}{a\sqrt{2}} = F(\omega).
\] (16.96)

16.6 The Dirac Delta Function and the Fourier transform

We recall the inverse Fourier transform

\[
f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} \, d\omega
\] (16.97)

so that substituting

\[
F(\omega) = \mathcal{F}[f(t)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t') e^{-i\omega t'} \, dt'
\] (16.98)
we have
\[ f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t') e^{-i\omega t'} dt' \right] e^{i\omega t} d\omega = \int_{-\infty}^{\infty} f(t') \left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega(t-t')} d\omega \right] dt' \]
(16.99)

Recalling the property of the Dirac Delta function
\[ f(t) = \int_{-\infty}^{\infty} f(t') \sigma(t-t') dt' \]
(16.100)
we find
\[ \sigma(t-t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega(t-t')} d\omega \]
(16.101)
The Dirac Delta function has different forms. One of this form is given by
\[
\sigma(t-t') = \lim_{a \to \infty} \int_{-a}^{a} e^{-i\omega(t-t')} d\omega = \frac{1}{2\pi} \lim_{a \to \infty} \frac{e^{i\omega(t-t')} - e^{-i\omega(t-t')}}{i(t-t')}
\]
\Rightarrow \quad \sigma(t-t') = \lim_{a \to \infty} \frac{\sin[a(t-t')]}{\pi(t-t')}.
(16.102)

Suppose the function
\[ f(t) = 1 \Rightarrow f(t') = 1 \]
(16.103)
for
\[ f(t) = \int_{-\infty}^{\infty} f(t') \sigma(t-t') dt' \]
(16.104)
we then find
\[ \int_{-\infty}^{\infty} \sigma(t-t') dt' = 1. \]
(16.105)
Now let’s examine the graph of the Dirac Delta function
\[ \sigma(t-t') = \lim_{a \to \infty} \frac{\sin[a(t-t')]}{\pi(t-t')}. \]
(16.106)
for different value of \( a \) at \( t' = 0 \).
From the above graphs it is easy to conclude that
\[ \sigma(t-t') = \begin{cases} \infty & t = t' \\ 0 & otherwise \end{cases} \]
(16.107)
which also leads to
\[ \int_{a}^{b} f(t') \sigma(t-t') dt' = \begin{cases} f(t) & a < t < b \\ 0 & otherwise \end{cases} \]
(16.108)
16.7 Applications of the Fourier Transform

The Time-Dependent Schrödinger Equation

\[- \frac{\hbar^2}{2m} \nabla^2 \Psi (\vec{r}, t) + U (\vec{r}) \Psi (\vec{r}, t) = i\hbar \frac{\partial \Psi (\vec{r}, t)}{\partial t} \]  

(16.109)

The Special Case of a Free Particle Traveling in the Positive-x Direction: For a free particle, we have

\[ U (\vec{r}) = 0 \]  

(16.110)

and the time independent part of the S.E. for a particle traveling in the positive x-direction can be written as

\[- \frac{\hbar^2}{2m} \frac{d^2 \psi (x)}{dx^2} = E \psi (x) \Rightarrow \frac{d^2 \psi (x)}{dx^2} = -\frac{2mE}{\hbar^2} \psi (x). \]  

(16.111)

Using De Broglie wavelength of the particle, we can relate the momentum with the wave vector

\[ \lambda = \frac{\hbar}{p} \Rightarrow p = \frac{\hbar}{\lambda} = \frac{2\pi \hbar}{2\pi} \Rightarrow p = k\hbar. \]  

(16.112)
Then the total energy for a free particle

$$E = \frac{p^2}{2m}$$  \hspace{1cm} (16.113)

can be expressed as

$$E = \frac{k^2 \hbar^2}{2m}.$$  \hspace{1cm} (16.114)

Using this result one finds

$$\frac{d^2 \psi (x)}{dx^2} = - \frac{2mE}{\hbar^2} \psi (x) = -k^2 \psi (x) \Rightarrow \frac{d^2 \psi (x)}{dx^2} + k^2 \psi (x) = 0$$  \hspace{1cm} (16.115)

the solution of which is given by

$$\psi (x) = A (k) e^{ikx}.$$  \hspace{1cm} (16.116)

Since for a free particle \( k \) takes continuous value, we may write the general solution as

$$\psi (x) = \int_{-\infty}^{\infty} A (k) e^{ikx} dk.$$  \hspace{1cm} (16.117)

We do not know anything about about \( A(k) \). What we do know, from a pure mathematical point of view, is a constant resulting from integrating a second order differential equation. From a physical point of view it is determined by quantum mechanical conditions that need to be met for a free particle. In quantum mechanics everything is probabilistic which is determined by the wave function. The probability of finding the particle anywhere \((-\infty < x < +\infty)\) is one. That means

$$\int_{-\infty}^{\infty} |\psi (x)|^2 dx = \int_{-\infty}^{\infty} \psi^* (x) \psi (x) dx = 1.$$  \hspace{1cm} (16.118)

Substituting the integral expression for \( \psi (x) \), we find

$$\int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} A^* (k') e^{-ik'x} dk' \right] \left[ \int_{-\infty}^{\infty} A (k) e^{ikx} dk \right] dx = 1$$

$$\Rightarrow \int_{-\infty}^{\infty} A (k) \int_{-\infty}^{\infty} A^* (k') \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix(k-k')} dx \right) dk' dk = \frac{1}{2\pi}.$$  \hspace{1cm} (16.119)

Applying the definition of the Dirac Delta function

$$\sigma (t-t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega(t-t')} d\omega,$$  \hspace{1cm} (16.120)
we may write

\[ \int_{-\infty}^{\infty} A^*(k') \sigma(k - k') dk' \] \[ = \frac{1}{2\pi} \tag{16.121} \]

and using the relation

\[ f(t) = \int_{-\infty}^{\infty} f(t') \sigma(t - t') dt' \] \[ \tag{16.122} \]

we find

\[ \int_{-\infty}^{\infty} A(k) A^*(k) \, dk = \frac{1}{2\pi} \Rightarrow \int_{-\infty}^{\infty} |A(k)|^2 \, dk = \frac{1}{2\pi}. \] \[ \tag{16.123} \]

Expressing \( A(k) \) as

\[ A(k) = \frac{1}{\sqrt{2\pi}} \phi(k), \] \[ \tag{16.124} \]

where \( \phi(k) \) is known as the momentum eigen function (which you will be introduce in quantum mechanics class), we find

\[ \int_{-\infty}^{\infty} |\phi(k)|^2 \, dk = \int_{-\infty}^{\infty} \phi^*(k) \phi(k) \, dk = 1. \] \[ \tag{16.125} \]

Then in terms of the momentum eigen function, the wave function can be expressed as

\[ \psi(x) = \int_{-\infty}^{\infty} A(k) e^{ikx} \, dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(k) e^{ikx} \, dk. \] \[ \tag{16.126} \]

This shows that the wave function for a free particle is the inverse Fourier Transform of the momentum eigen function \( \phi(k) \)

**Uncertainties**

\[ \Delta x = \sigma_x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}, \Delta k = \sigma_k = \sqrt{\langle k^2 \rangle - \langle k \rangle^2} \] \[ \tag{16.127} \]

**Expectation Values**

\[ \langle f(x) \rangle = \int_{-\infty}^{\infty} \psi^*(x) f(x) \psi(x) \, dx, \langle g(k) \rangle = \int_{-\infty}^{\infty} \phi^*(k) g(k) \phi(k) \, dk \] \[ \tag{16.128} \]
Example 15.9 For a particle described by a Gaussian Wavepacket

\[ \psi(x) = A e^{-\alpha x^2} = \left( \frac{2\alpha}{\pi} \right)^{1/4} e^{-\alpha x^2} \]  

(16.129)

Some Useful Integrals

\[ \int_{-\infty}^{\infty} e^{-ay^2} dy = \sqrt{\frac{\pi}{a}}, \int_{-\infty}^{\infty} y^2 e^{-y^2} dy = \frac{\sqrt{\pi}}{2} \]  

(16.130)

(a) Show that the momentum eigen function is given by

\[ \phi(k) = \frac{A}{\sqrt{2\alpha}} e^{-k^2/4\alpha} \]  

(16.131)

(b) Show that the expectation values

\[ \langle k \rangle = 0, \langle k^2 \rangle = \alpha. \]  

(16.132)

(c) Find the Expectation values \( \langle x \rangle \) and \( \langle x^2 \rangle \).

(d) Show that the Heisenberg Uncertainty Principle

\[ \Delta x \Delta p_x \geq \frac{\hbar}{2} \]  

(16.133)

is satisfied.

Solution:

(a) We recall

\[ \psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(k) e^{ikx} dk \]  

(16.134)

which is the inverse Fourier transform of the momentum eigen function \( \phi(k) \), we can write

\[ \phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x) e^{-ikx} dx. \]  

(16.135)

Then substituting

\[ \psi(x) = A e^{-\alpha x^2} = \left( \frac{2\alpha}{\pi} \right)^{1/4} e^{-\alpha x^2} \]  

(16.136)
we find
\[ \phi(k) = \frac{A}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ax^2} e^{-ikx} dx = \frac{A}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-a(x^2 + \frac{ikx}{\alpha})} dx. \] (16.137)

Noting that
\[ x^2 + \frac{ik}{\alpha} = \left(x + \frac{ik}{2\alpha}\right)^2 + \frac{k^2}{4\alpha^2} \] (16.138)
we can write
\[ \phi(k) = \frac{A}{\sqrt{2\pi}} e^{-\frac{k^2}{4\alpha^2}} \int_{-\infty}^{\infty} e^{-a(x + \frac{ik}{2\alpha})^2} dx. \] (16.139)

Using the result in the previous example or Mathematica

\[ \text{In[3]} = \int_{-\infty}^{\infty} \exp[-ax^2] \, dx \]

\[ \text{Out[3]}= \text{If}[\text{Re}[a] > 0, \frac{\sqrt{\pi}}{\sqrt{a}}, \text{Integrate}[e^{-ax^2}, \{x, -\infty, \infty\}, \text{Assumptions} \to \text{Re}[a] \neq 0]] \]

we can easily show that
\[ \phi(k) = \frac{A}{\sqrt{2\alpha}} e^{-\frac{k^2}{4\pi}}. \] (16.140)

(b) In the momentum space we can express the expectation values
\[ \langle k \rangle = \int_{-\infty}^{\infty} \phi^*(k) k \phi(k) \, dk, \langle k^2 \rangle = \int_{-\infty}^{\infty} \phi^*(k) k^2 \phi(k) \, dk \] (16.141)
and using the momentum eigen function, we may write
\[ \langle k \rangle = \left(\frac{A}{\sqrt{2\alpha}}\right)^2 \int_{-\infty}^{\infty} k e^{-\frac{k^2}{4\pi}} \, dk = 0, \langle k^2 \rangle = \left(\frac{A}{\sqrt{2\alpha}}\right)^2 \int_{-\infty}^{\infty} k^2 e^{-\frac{k^2}{4\pi}} \, dk. \] (16.142)

Using the transformation of variable
\[ x = \frac{k}{\sqrt{2\alpha}} \Rightarrow dk = \sqrt{2\alpha} \, dx \] (16.143)
we have
\[ \langle k^2 \rangle = \left(\frac{A}{\sqrt{2\alpha}}\right)^2 \int_{-\infty}^{\infty} 2\alpha x^2 e^{-x^2} \sqrt{2\alpha} \, dx = \left(\frac{A}{\sqrt{2\alpha}}\right)^2 (2\alpha)^{3/2} \int_{-\infty}^{\infty} x^2 e^{-x^2} \, dx. \] (16.144)
so that using the relation
\[ \int_{-\infty}^{\infty} y^2 e^{-y^2} \, dy = \frac{\sqrt{\pi}}{2} \]  
(16.145)
we find
\[ \langle k^2 \rangle = \left( \frac{A}{\sqrt{2\alpha}} \right)^2 (2\alpha)^{3/2} \frac{\sqrt{\pi}}{2} = A^2 (2\alpha)^{1/2} \frac{\sqrt{\pi}}{2}. \]  
(16.146)
Substituting
\[ A = \left( \frac{2\alpha}{\pi} \right)^{1/4} \]  
(16.147)
we find
\[ \langle k^2 \rangle = \frac{\sqrt{2\alpha}}{\sqrt{\pi}} \left( \frac{\sqrt{2\alpha}}{2} \right) \frac{\sqrt{\pi}}{2} = \alpha. \]  
(16.148)

(c) It is easier to find these expectation values in position space rather than in momentum space, which we may write as
\[ \langle x \rangle = \int_{-\infty}^{\infty} \psi^* (x) x \psi (x) \, dx, \quad \langle x^2 \rangle = \int_{-\infty}^{\infty} \psi^* (x) x^2 \psi (x) \, dx. \]  
(16.149)
Using the wave function
\[ \psi (x) = A e^{-\alpha x^2} = \left( \frac{2\alpha}{\pi} \right)^{1/4} e^{-\alpha x^2} \]  
(16.150)
the expectation values become
\[ \langle x \rangle = A^2 \int_{-\infty}^{\infty} x e^{-2\alpha x^2} \, dx = 0, \quad \langle x^2 \rangle = A^2 \int_{-\infty}^{\infty} x^2 e^{-2\alpha x^2} \, dx. \]  
(16.151)
Replacing
\[ y = \sqrt{2\alpha} x \Rightarrow dx = \frac{dy}{\sqrt{2\alpha}} \]  
(16.152)
we have
\[ \langle x^2 \rangle = \frac{A^2}{(\sqrt{2\alpha})^{3/2}} \int_{-\infty}^{\infty} y^2 e^{-y^2} \, dy \]  
(16.153)
which gives
\[ \langle x^2 \rangle = \frac{A^2}{(\sqrt{2\alpha})^{3/2}} \frac{\sqrt{\pi}}{2} \]  
(16.154)
or
\[ \langle x^2 \rangle = \frac{1}{(\sqrt{2\alpha})^{3/2}} \frac{\sqrt{2\alpha} \sqrt{\pi}}{2} = \frac{1}{4\alpha} \]  
(16.155)
(d) The Uncertainty in position and momentum

$$\Delta x = \sigma_x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}, \Delta p_x = \hbar \Delta k = \hbar \sigma_k = \hbar \sqrt{\langle k^2 \rangle - \langle k \rangle^2}$$  \hspace{1cm} (16.156)$$

would then be

$$\Delta x = \sigma_x = \frac{1}{2\sqrt{\alpha}}, \Delta p_x = \hbar \Delta k = \hbar \sigma_k = \hbar \sqrt{\alpha}.$$ \hspace{1cm} (16.157)$$

Then the uncertainty principle

$$\Delta x \Delta p_x = \frac{1}{2\sqrt{\alpha}} \hbar \sqrt{\alpha} = \frac{\hbar}{2}$$ \hspace{1cm} (16.158)$$

is satisfied. This means the Gaussian wave packet represent the minimum uncertainty state.

### 16.8 Fourier Transforms and Convolution

**Convolution:** For two functions $f(x)$ and $g(x)$ with corresponding Fourier Transforms, $F(\omega)$ and $G(\omega)$, respectively, the convolution of these two functions over the interval $(-\infty, \infty)$, denoted by $f \ast g$, is defined as

$$f \ast g = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(y) f(x-y) \, dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y) g(x-y) \, dy = g \ast f.$$ \hspace{1cm} (16.159)$$

This form of integrals appears in probability theory in the determination of the probability density of two random, independent variables. One good example is the solution of the Poisson’s equation

$$V(\vec{r}) = \frac{1}{4\pi \varepsilon_0} \int \frac{\rho(\vec{r'})}{|\vec{r} - \vec{r'}|} \, d^3r'.$$ \hspace{1cm} (16.160)$$

This may be interpreted as convolution of a charge distribution $\rho(\vec{r'})$ (as $g(y)$) and a weighting function $(4\pi \varepsilon_0 |\vec{r} - \vec{r'}|)^{-1}$ (as $f(x-y)$). The operation of convolution sometimes referred to as the Faltung, the German term for "folding".

**The Convolution theorem:** the convolution theorem states that the convolution of two or more functions is the same as the Fourier inverse transforms of a product of Fourier Transforms. That means

$$f_1 \ast f_2 \ast f_3 \ldots f_n = \mathcal{F}^{-1} [F_1(\omega) F_2(\omega) F_3(\omega) \ldots F_n(\omega)]$$ \hspace{1cm} (16.161)$$

**Proof:** We shall proof this theorem by considering the case of two functions

$$f_1 \ast f_2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f_2(y) f_1(x-y) \, dy$$ \hspace{1cm} (16.162)$$

Let Fourier Transform of the function $f_1(x-y)$ is $F_1(\omega)$, then may write

$$f_1(x-y) = \mathcal{F}^{-1} [F_1(\omega)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_1(\omega) e^{i\omega(x-y)} \, d\omega$$ \hspace{1cm} (16.163)$$
so that the convolution integral can be put in the form

\[
\begin{aligned}
    f_1 * f_2 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f_1(y) \frac{1}{\sqrt{2\pi}} \left( \int_{-\infty}^{\infty} F_1(\omega) e^{i\omega(y-x)} \, d\omega \right) \, dy \\
    f_1 * f_2 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_1(\omega) \frac{1}{\sqrt{2\pi}} \left( \int_{-\infty}^{\infty} f_2(y) e^{-i\omega y} \, dy \right) e^{i\omega x} \, d\omega
\end{aligned}
\]

Noting that

\[
    F_2(\omega) = \int_{-\infty}^{\infty} f_2(y) e^{-i\omega y} \, dy
\]

we can write the above expression as

\[
    f_1 * f_2 = \frac{1}{\sqrt{2\pi}} \left[ \int_{-\infty}^{\infty} F_1(\omega) F_2(\omega) e^{i\omega x} \, d\omega \right] = \mathcal{F}^{-1} [F_1(\omega) F_2(\omega)]
\]

**Example 15.10** Discuss, compute, and plot the convolution of the functions \(f(x)\) and \(g(x)\) given below.

**Solution:** We note that the two functions shown above can be expressed as

\[
    f(x) = \begin{cases} 
    1 & -10 < x < 10 \\
    0 & \text{otherwise}
    \end{cases}
\]

(16.168)

\[
    g(x) = \begin{cases} 
    \frac{1}{10} x & 0 < x < 10 \\
    0 & \text{otherwise}
    \end{cases}
\]

(16.169)

For \(g(x-y)\), we may write

\[
    g(x-y) = \begin{cases} 
    \frac{1}{10} (x-y) & 0 < x-y < 10 \\
    0 & \text{otherwise}
    \end{cases}
\]

(16.170)

or

\[
    g(x-y) = \begin{cases} 
    \frac{1}{10} (x-y) & x-10 < y < x \\
    0 & \text{otherwise}
    \end{cases}
\]

(16.171)

Applying the convolution formula

\[
    f * g = \frac{1}{\sqrt{2\pi}} \int_{x-10}^{x} f(y) \frac{1}{10} (x-y) \, dy = \frac{1}{10\sqrt{2\pi}} \int_{x-10}^{x} f(y) (x-y) \, dy
\]

(16.172)

this can be rewritten as

\[
    f * g = \frac{1}{10\sqrt{2\pi}} \int_{x-10}^{x} f(y) (x-y) \, dy
\]

(16.173)

we have to be careful in evaluating this integral because of the function \(f(y)\).
Case 1: \( x \leq -10 \)

For this case the range of \( y \) would be \((-\infty, -20)\). In this interval we have \( f(y) = 0 \), and we find

\[
f * g = 0
\]

\[\text{(16.174)}\]

Case 2: \(-10 < x < 0\)

For this case the range of \( y \) would be \((-20, 0)\). In this interval we have \( f(y) = 0 \) for \( y < -10 \) and \( f(y) = 1 \) for \(-10 < y < 0 \), and we find

\[
f * g = \frac{1}{10\sqrt{2\pi}} \int_{-10}^{x} (x-y) \, dy = \frac{1}{10\sqrt{2\pi}} \left[ \frac{xy-y^2}{2} \right]_{-10}^{x}
\]

\[= \frac{1}{10\sqrt{2\pi}} \left( \frac{x^2}{2} + 10x + 50 \right) \Rightarrow f * g = \frac{1}{\sqrt{2\pi}} \left( \frac{x^2}{20} + x + 5 \right)
\]

\[\text{(16.175)}\]

Case 3: \( 0 < x < 10 \)

For this case the range of \( y \) would be \((-10, 10)\). In this interval we have \( f(y) = 1 \), and we find

\[
f * g = \frac{1}{10\sqrt{2\pi}} \int_{-10}^{x} (x-y) \, dy = \frac{1}{10\sqrt{2\pi}} \left[ \frac{xy-y^2}{2} \right]_{-10}^{x}
\]

\[= \frac{1}{10\sqrt{2\pi}} \left( \frac{x^2}{2} - x \left( (x-10) - \frac{(x-10)^2}{2} \right) \right)
\]

\[= \frac{1}{10\sqrt{2\pi}} \left( \frac{x^2}{2} - \left( \frac{x^2 - 100}{2} \right) \right) \Rightarrow f * g = \frac{5}{\sqrt{2\pi}}
\]

\[\text{(16.176)}\]

Case 4: \( 10 < x < 20 \)

For this case the range of \( y \) would be \((0, 20)\). In this interval we have \( f(y) = 0 \) for \( 10 < y < 20 \) and \( f(y) = 1 \) for \( 0 < y < 10 \), and we find

\[
f * g = \frac{1}{10\sqrt{2\pi}} \int_{-10}^{10} (x-y) \, dy = \frac{1}{10\sqrt{2\pi}} \left[ \frac{xy-y^2}{2} \right]_{-10}^{10}
\]

\[= \frac{1}{10\sqrt{2\pi}} \left( \frac{20x - 100 - x^2 + 100}{2} \right)
\]

\[\Rightarrow f * g = \frac{1}{\sqrt{2\pi}} \left( x - \frac{x^2}{20} \right)
\]

\[\text{(16.177)}\]

The graph of the convoluted functions is shown below
Important Properties of Convolution

(a) \( f(x) * g(x) = g(x) * f(x) \)

(b) convolution tends to smooth out \( f(x) \)

(c) convolution tends to spread out \( f(x) \)
Part III

Mathematical Methods in physics III
Chapter 17

Manifolds

17.1 What is a Manifold?

Consider a ridged meterstick pinned at the north and south poles inside a hollow sphere as shown in Fig.17.1. Initially, $t = 0$, the sphere is at rest and suddenly begins to rotate. It is free to rotate about the x-axis, y-axis, or z-axis. Let’s say we want to describe the angular position of the center of mass of the meter stick over a period of time, $\tau = 10s$, with a time interval of 2s. How many independent parameters, that depend on time, do we need to describe the angular position of the center of mass of the meterstick relative to its initial position at $t = 0$? Well the answer is simple. We need three independent parameters, the Euler
angles, \((\alpha_1(t), \alpha_2(t), \alpha_3(t))\) which describes the rotation about the \(x, y,\) and \(z\) axes at a given instant of time. Then over the 10 second interval we have a
set that consist of 5 points
\[
\{ [\alpha_1(2), \alpha_2(2), \alpha_3(2)], [\alpha_1(4), \alpha_2(4), \alpha_3(4)], [\alpha_1(6), \alpha_2(6), \alpha_3(6)], \\
[\alpha_1(8), \alpha_2(8), \alpha_3(8)], [\alpha_1(10), \alpha_2(10), \alpha_3(10)] \}
\]
We can make the time interval infinitesimal to continuously describe the angular position of the center of mass of the meterstick. The resulting set of points form a Manifold of dimension three.

Let’s consider another example of a Manifold. In classical mechanics you may have studied what is called the phase space. In this space you can describe the state of a particle over a period of time using the three coordinates of space locating the position of the particle and the three coordinates of speed (or momentum) describing how fast the particle is moving at a given instant of time. In Cartesian coordinate system we use the independent parameters \((x(t), y(t), z(t))\) for position and \((\dot{x}(t), \dot{y}(t), \dot{z}(t))\) for the speed of the particle. I can represent these independent parameters by \((x^1(t), x^2(t), x^3(t), x^4(t), x^5(t), x^6(t))\). So when you describe the state of the particle say from \(t = 0\) to \(t = t_0\), you can use infinitesimal time interval so that you will have a set of points that can be parameterized continuously in terms of \((x^1(t), x^2(t), \ldots, x^6(t))\). These set of points form a Manifold of dimension six.

Now let’s apply this to the Minkowski spacetime where we have three space coordinates \((x(t), y(t), z(t))\) and time, \(t\). This forms a four dimensional manifol with four coordinates each parametrized by the proper time \(\tau\), \((x^1(\tau), x^2(\tau), x^3(\tau), x^4(\tau))\). Therefore, a manifold is any set that can be continuously parameterized. Therefore, an \(N\) dimensional manifold, \(M\), of points is one for which \(N\) independent real coordinates \((x^1, x^2, x^3, \ldots, x^N)\) are required to specify any point completely.

A manifold is Continuous: if you pick any point, \(p\), on the Manifold and you can find another points whose coordinates differ infinitesimally from the point \(p\).

A manifold is differentiable: if you pick any point, \(p\), on the Manifold and you can find a scalar function \(\phi\) that is differentiable at that point \(p\).

Coordinates of a Manifold \(M\): a point in an \(N\)-Dimensional Manifold is represented by the coordinates \((x^1, x^2, x^3, \ldots, x^N)\) which we represent by \(x^a\) where it is understood that \(a = 1, 2, 3 \ldots, N\).

Degeneracy in a Manifold: sometimes it may not be possible to cover the whole manifold with only one none-degenerate coordinate system. Example is a plane in polar coordinate system \((\rho, \varphi)\). A plane is a two dimension Manifold. (called \(R^2\)). A plane in polar coordinates has a degeneracy at the origin since \(\varphi\) is indeterminate at the origin. (Fig. 17.2)

Coordinate patches: these are coordinate systems that covers a portion of the Manifold where we have degeneracy. For example the surface of a sphere is a two dimensional Manifold (called \(S^2\)). It can be described by two independent coordinates \((x^1 = \theta, x^2 = \varphi)\) except at two points on the Manifold. These are
17.2 CURVES AND SURFACES IN A MANIFOLD

Figure 17.2: The degenerate point on a plane in polar coordinates.

Figure 17.3: The north and south pole on the surface of sphere are degenerate in polar coordinates.

the north and south pole where $\varphi$ is indeterminate (or there is degeneracy). (Fig.17.3) There is no coordinate system that covers the entire sphere without running into these two degenerate points. In this case the smallest number of patches we need is two.

*Atlas:* an atlas is a set of coordinate patches that covers the whole Manifold.

### 17.2 Curves and surfaces in a Manifold

Both curves and surfaces on a Manifold are defined *parametrically*. That means we use some common parameters. For example, a curve in the phase space that we saw earlier, can be defined in terms of the time parameter, $t$. Another example, a curve in the 4D Minkowski spacetime manifold is defined by the interval

$$ds^2 (\tau) = c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2.$$  \hspace{1cm} (17.1)
Generally, we use a parameter, \( u \), to define a curve.

A curve: a curve in a Manifold of dimension \( N \) is defined by a parametric equation

\[
x^a = x^a(u), \text{ where } a = 1, 2, 3...N.
\]  

(17.2)

For example, the curve shown in pink in Fig. 17.4 is defined by

\[
\begin{align*}
x^1(u) &= \cos^5(u), \\
x^2(u) &= 0.4u^2, \\
x^3(u) &= 0.4u^3,
\end{align*}
\]  

(17.3)

and it needs only one parameter, \( u \).

A surface: a surface in a Manifold of dimension \( N \) (which also referred as a submanifold) has \( M \) degrees of freedom which is always less than the dimension of the Manifold (\( M < N \)) and therefore it depends on \( M \) parameters that we represent by \((u^1, u^2, u^3, ... u^M)\) and is defined by the parametric equation

\[
x^a = x^a(u^1, u^2, u^3, ... u^M), \text{ where } a = 1, 2, 3...N.
\]  

(17.4)

Hypersurface: a surface of dimension \( M \) in a Manifold of dimension \( N \) with \( M = N - 1 \). In this case, the \( N - 1 \) parameters can be eliminated from the \( N \) equations and you can find one equation

\[
f(x^1, x^2, x^3, ... x^N) = 0.
\]  

(17.5)

The surface shown in blue in Fig. 17.4 in the 3D manifold needs two parameters to define it

\[
x^1 = 2\cos(u_1), \quad x^2 = \sin(u_2), \quad x^3 = u_2.
\]  

(17.6)

Note that this surface is in 3D manifold and it parameterized by two coordinates \((u, v)\), \((M = N - 1 = 3 - 1 = 2\), it is a hypersurface).
Example 2.1 Let’s consider the 3-D Euclidean Manifold. A sphere is a hypersurface since $M = 2$, (Fig.17.5). A point on a sphere is defined by
\[ x^2 + y^2 + z^2 = a^2, \]  
(17.7)
where $a$ is the radius of the sphere. We note that in this case the surface of the sphere is a hypersurface that can be defined by the equation
\[ g(x^1, x^2, x^3) = (x^1)^2 + (x^2)^2 + (x^3)^2 - a^2 = 0, \]  
(17.8)
where we used $(x^1, x^2, x^3)$ for $(x, y, z)$. Introducing the parameters $u_1, u_2,$ and $u_3$ defined by
\[ x^1 = u_1 \sin(u_2) \cos(u_3), \quad x^2 = u_1 \sin(u_2) \sin(u_3), \quad x^3 = u_1 \cos(u_2) \]
we can write the equation that define the surface of the sphere ($M = 2$) that is embedded in a 3D manifold ($N = 3$) using only one parameter (by eliminating the $N - 1 = 2$, parameters) as
\[ g(x^1, x^2, x^3) = (u_1 \sin(u_2) \cos(u_3))^2 + (u_1 \sin(u_2) \sin(u_3))^2 + (u_1 \cos(u_2))^2 - a^2 = 0 \Rightarrow g(x^1, x^2, x^3) = u_1^2 - a^2 = 0 \]  
(17.9)
which is the property of a hypersurface in a manifold.

Therefore a point is restricted to lie in a hypersurface ($M = N - 1$ dimensional submanifold embedded in $N$-dimensional Manifold), then the points coordinate must satisfy Eq. (17.5). We come up with a similar generalization to this for a point that belong to any surface with dimension $M$ in a Manifold.
CHAPTER 17. MANIFOLDS

of dimension $N$ ($M < N$).

$$
g_1 (x^1, x^2, x^3, \ldots, x^N) = 0, \quad (17.10)$$
$$
g_2 (x^1, x^2, x^3, \ldots, x^N) = 0, \quad (17.11)$$
$$
g_3 (x^1, x^2, x^3, \ldots, x^N) = 0$$
$$\quad \ldots$$
$$\quad \ldots$$
$$
g_{N-M} (x^1, x^2, x^3, \ldots, x^N) = 0
$$

17.3 Coordinate transformations and summation convention

Let's consider the 3-D Euclidean Manifold. A point in this Manifold can be represented using Cartesian coordinates $(x, y, z)$ which we shall represent by $(x^1, x^2, x^3)$. This same point can also be represented using spherical coordinates $(r, \theta, \phi)$ that shall represent by $(x^1, x^2, x^3)$. Now the question is how we relate the Cartesian coordinates with the spherical coordinates or vice versa. In terms of these notations, one can write

$$
r \to r (x, y, z), \quad \text{or} \quad x^1 \to x^1 (x^1, x^2, x^3), \quad (17.11)$$
$$\theta \to \theta (x, y, z), \quad \text{or} \quad x^2 \to x^2 (x^1, x^2, x^3),$$
$$\phi \to \phi (x, y, z), \quad \text{or} \quad x^3 \to x^3 (x^1, x^2, x^3),$$

or

$$
x \to x (r, \theta, \phi), \quad \text{or} \quad x^1 \to x^1 (x^1, x^2, x^3), \quad (17.12)$$
$$y \to y (r, \theta, \phi), \quad \text{or} \quad x^2 \to x^2 (x^1, x^2, x^3),$$
$$z \to z (r, \theta, \phi), \quad \text{or} \quad x^3 \to x^3 (x^1, x^2, x^3).$$

Suppose we have a function, $g (x, y, z)$ or $g (r, \theta, \phi)$, which can be expressed by $g (x^1, x^2, x^3)$ and $g (x^1, x^2, x^3)$, respectively, then one can write [Theoretical Physics I],

$$
\frac{\partial g}{\partial x^1} = \frac{\partial g}{\partial x^1} \frac{\partial x^1}{\partial x^1} + \frac{\partial g}{\partial x^2} \frac{\partial x^2}{\partial x^1} + \frac{\partial g}{\partial x^3} \frac{\partial x^3}{\partial x^1}, \quad (17.13)$$
$$
\frac{\partial g}{\partial x^2} = \frac{\partial g}{\partial x^1} \frac{\partial x^2}{\partial x^1} + \frac{\partial g}{\partial x^2} \frac{\partial x^2}{\partial x^2} + \frac{\partial g}{\partial x^3} \frac{\partial x^3}{\partial x^2}, \quad (17.14)$$
$$
\frac{\partial g}{\partial x^3} = \frac{\partial g}{\partial x^1} \frac{\partial x^3}{\partial x^1} + \frac{\partial g}{\partial x^2} \frac{\partial x^2}{\partial x^3} + \frac{\partial g}{\partial x^3} \frac{\partial x^3}{\partial x^3}. \quad (17.15)$$

Using matrices [Theoretical Physics I] this can be put in the form

$$
\begin{bmatrix}
\frac{\partial g}{\partial x^1} \\
\frac{\partial g}{\partial x^2} \\
\frac{\partial g}{\partial x^3}
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^2}{\partial x^1} & \frac{\partial x^3}{\partial x^1} \\
\frac{\partial x^1}{\partial x^2} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^3}{\partial x^2} \\
\frac{\partial x^1}{\partial x^3} & \frac{\partial x^2}{\partial x^3} & \frac{\partial x^3}{\partial x^3}
\end{bmatrix}
\begin{bmatrix}
\frac{\partial g}{\partial x^1} \\
\frac{\partial g}{\partial x^2} \\
\frac{\partial g}{\partial x^3}
\end{bmatrix} \quad . \quad (17.16)
$$
17.3. COORDINATE TRANSFORMATIONS AND SUMMATION CONVENTION

For the inverse case, following a similar procedure, we can write

\[
\begin{bmatrix}
\frac{\partial q}{\partial x^1} & \frac{\partial q}{\partial x^2} & \frac{\partial q}{\partial x^3} \\
\frac{\partial q}{\partial x^4} & \frac{\partial q}{\partial x^5} & \frac{\partial q}{\partial x^6} \\
\frac{\partial q}{\partial x^7} & \frac{\partial q}{\partial x^8} & \frac{\partial q}{\partial x^9}
\end{bmatrix} = \begin{bmatrix}
\frac{\partial x^1}{\partial g^1} & \frac{\partial x^2}{\partial g^1} & \frac{\partial x^3}{\partial g^1} \\
\frac{\partial x^4}{\partial g^2} & \frac{\partial x^5}{\partial g^2} & \frac{\partial x^6}{\partial g^2} \\
\frac{\partial x^7}{\partial g^3} & \frac{\partial x^8}{\partial g^3} & \frac{\partial x^9}{\partial g^3}
\end{bmatrix}
\]

so that

\[
\begin{bmatrix}
\frac{\partial g}{\partial x^1} & \frac{\partial g}{\partial x^2} & \frac{\partial g}{\partial x^3} \\
\frac{\partial g}{\partial x^4} & \frac{\partial g}{\partial x^5} & \frac{\partial g}{\partial x^6} \\
\frac{\partial g}{\partial x^7} & \frac{\partial g}{\partial x^8} & \frac{\partial g}{\partial x^9}
\end{bmatrix} = \begin{bmatrix}
\frac{\partial x^1}{\partial g^1} & \frac{\partial x^2}{\partial g^1} & \frac{\partial x^3}{\partial g^1} \\
\frac{\partial x^4}{\partial g^2} & \frac{\partial x^5}{\partial g^2} & \frac{\partial x^6}{\partial g^2} \\
\frac{\partial x^7}{\partial g^3} & \frac{\partial x^8}{\partial g^3} & \frac{\partial x^9}{\partial g^3}
\end{bmatrix}
\]

(17.17)

can be written as

\[
\begin{bmatrix}
\frac{\partial g}{\partial x^1} & \frac{\partial g}{\partial x^2} & \frac{\partial g}{\partial x^3} \\
\frac{\partial g}{\partial x^4} & \frac{\partial g}{\partial x^5} & \frac{\partial g}{\partial x^6} \\
\frac{\partial g}{\partial x^7} & \frac{\partial g}{\partial x^8} & \frac{\partial g}{\partial x^9}
\end{bmatrix} = \begin{bmatrix}
\frac{\partial x^1}{\partial g^1} & \frac{\partial x^2}{\partial g^1} & \frac{\partial x^3}{\partial g^1} \\
\frac{\partial x^4}{\partial g^2} & \frac{\partial x^5}{\partial g^2} & \frac{\partial x^6}{\partial g^2} \\
\frac{\partial x^7}{\partial g^3} & \frac{\partial x^8}{\partial g^3} & \frac{\partial x^9}{\partial g^3}
\end{bmatrix} \begin{bmatrix}
\frac{\partial g}{\partial x^1} & \frac{\partial g}{\partial x^2} & \frac{\partial g}{\partial x^3} \\
\frac{\partial g}{\partial x^4} & \frac{\partial g}{\partial x^5} & \frac{\partial g}{\partial x^6} \\
\frac{\partial g}{\partial x^7} & \frac{\partial g}{\partial x^8} & \frac{\partial g}{\partial x^9}
\end{bmatrix}.
\]

(17.19)

There follows that

\[
\begin{bmatrix}
\frac{\partial x^1}{\partial g^1} & \frac{\partial x^2}{\partial g^1} & \frac{\partial x^3}{\partial g^1} \\
\frac{\partial x^4}{\partial g^2} & \frac{\partial x^5}{\partial g^2} & \frac{\partial x^6}{\partial g^2} \\
\frac{\partial x^7}{\partial g^3} & \frac{\partial x^8}{\partial g^3} & \frac{\partial x^9}{\partial g^3}
\end{bmatrix} = 1.
\]

(17.20)

This means the matrix

\[
A^{-1} = \begin{bmatrix}
\frac{\partial x^1}{\partial g^1} & \frac{\partial x^2}{\partial g^1} & \frac{\partial x^3}{\partial g^1} \\
\frac{\partial x^1}{\partial g^2} & \frac{\partial x^2}{\partial g^2} & \frac{\partial x^3}{\partial g^2} \\
\frac{\partial x^1}{\partial g^3} & \frac{\partial x^2}{\partial g^3} & \frac{\partial x^3}{\partial g^3}
\end{bmatrix}
\]

(17.21)

must be the inverse matrix for

\[
A = \begin{bmatrix}
\frac{\partial x^1}{\partial g^1} & \frac{\partial x^2}{\partial g^1} & \frac{\partial x^3}{\partial g^1} \\
\frac{\partial x^1}{\partial g^2} & \frac{\partial x^2}{\partial g^2} & \frac{\partial x^3}{\partial g^2} \\
\frac{\partial x^1}{\partial g^3} & \frac{\partial x^2}{\partial g^3} & \frac{\partial x^3}{\partial g^3}
\end{bmatrix}
\]

(17.22)

so that

\[
A^{-1}A = AA^{-1} = 1.
\]

(17.23)

We note that the transpose, \(A^T\), is given by

\[
A^T = \begin{bmatrix}
\frac{\partial x^1}{\partial x^1} & \frac{\partial x^1}{\partial x^2} & \frac{\partial x^1}{\partial x^3} \\
\frac{\partial x^2}{\partial x^1} & \frac{\partial x^2}{\partial x^2} & \frac{\partial x^2}{\partial x^3} \\
\frac{\partial x^3}{\partial x^1} & \frac{\partial x^3}{\partial x^2} & \frac{\partial x^3}{\partial x^3}
\end{bmatrix}
\]

(17.24)

Similarly for the inverse matrix, the transpose matrix which we express as

\[
\frac{\partial x'^\alpha}{\partial x^\beta} = \begin{bmatrix}
\frac{\partial x'^1}{\partial x^1} & \frac{\partial x'^1}{\partial x^2} & \frac{\partial x'^1}{\partial x^3} \\
\frac{\partial x'^2}{\partial x^1} & \frac{\partial x'^2}{\partial x^2} & \frac{\partial x'^2}{\partial x^3} \\
\frac{\partial x'^3}{\partial x^1} & \frac{\partial x'^3}{\partial x^2} & \frac{\partial x'^3}{\partial x^3}
\end{bmatrix}
\]

(17.25)
is the transformation matrix that transforms the coordinates \((x^1, x^2, x^3)\) to \((x'^1, x'^2, x'^3)\).

For a Manifold of dimension \(N\), the transformation matrix is given by

\[
\frac{\partial x'^a}{\partial x^b} = \begin{bmatrix}
\frac{\partial x'^1}{\partial x^1} & \frac{\partial x'^1}{\partial x^2} & \cdots & \frac{\partial x'^1}{\partial x^N} \\
\frac{\partial x'^2}{\partial x^1} & \frac{\partial x'^2}{\partial x^2} & \cdots & \frac{\partial x'^2}{\partial x^N} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial x'^N}{\partial x^1} & \frac{\partial x'^N}{\partial x^2} & \cdots & \frac{\partial x'^N}{\partial x^N}
\end{bmatrix}.
\]  

(17.26)

As we recall from [Theoretical physics I], a Matrix is invertible provided the determinant is different from zero. Therefore, the inverse transformation is possible provided the determinant of the transformation matrix which is known as the Jacobian, \(J\), is different from zero

\[
J = \det \left[ \frac{\partial x'^a}{\partial x^b} \right] = \begin{vmatrix}
\frac{\partial x'^1}{\partial x^1} & \frac{\partial x'^1}{\partial x^2} & \cdots & \frac{\partial x'^1}{\partial x^N} \\
\frac{\partial x'^2}{\partial x^1} & \frac{\partial x'^2}{\partial x^2} & \cdots & \frac{\partial x'^2}{\partial x^N} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial x'^N}{\partial x^1} & \frac{\partial x'^N}{\partial x^2} & \cdots & \frac{\partial x'^N}{\partial x^N}
\end{vmatrix} \neq 0
\]

(17.27)

The inverse transformation Matrix can be written as

\[
\frac{\partial x^a}{\partial x'^b} = \begin{bmatrix}
\frac{\partial x^1}{\partial x'^1} & \frac{\partial x^1}{\partial x'^2} & \cdots & \frac{\partial x^1}{\partial x'^N} \\
\frac{\partial x^2}{\partial x'^1} & \frac{\partial x^2}{\partial x'^2} & \cdots & \frac{\partial x^2}{\partial x'^N} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial x^N}{\partial x'^1} & \frac{\partial x^N}{\partial x'^2} & \cdots & \frac{\partial x^N}{\partial x'^N}
\end{bmatrix}
\]

(17.28)

and the Jacobian \(J'\)

\[
J' = \det \left[ \frac{\partial x^a}{\partial x'^b} \right] = \begin{vmatrix}
\frac{\partial x^1}{\partial x'^1} & \frac{\partial x^1}{\partial x'^2} & \cdots & \frac{\partial x^1}{\partial x'^N} \\
\frac{\partial x^2}{\partial x'^1} & \frac{\partial x^2}{\partial x'^2} & \cdots & \frac{\partial x^2}{\partial x'^N} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial x^N}{\partial x'^1} & \frac{\partial x^N}{\partial x'^2} & \cdots & \frac{\partial x^N}{\partial x'^N}
\end{vmatrix}
\]

(17.29)

We note that

\[
\frac{\partial x'^a}{\partial x^I} = \frac{\partial x^a}{\partial x^1} \frac{\partial x^1}{\partial x'^a} + \frac{\partial x^a}{\partial x^2} \frac{\partial x^2}{\partial x'^a} + \cdots + \frac{\partial x^a}{\partial x^N} \frac{\partial x^N}{\partial x'^a} = \sum_{b=1}^{N} \frac{\partial x^b}{\partial x'^a} \frac{\partial x^b}{\partial x^I}
\]

(17.30)
Noting that for independent coordinates
\[ \frac{\partial x^a}{\partial x^c} = \begin{cases} 
0, & a \neq c \\
1, & a = c 
\end{cases} \]  
we can generally write
\[ \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} \frac{\partial x^b}{\partial x^c} = \delta_c^a. \]  
(17.32)

Consider two points \( P \) and \( Q \) on a Manifold with dimension \( N \). Suppose these points are separated by infinitesimal interval so that if the coordinates of \( P \) is \( x^a \) and that of \( Q \) is \( x^a + dx^a \), then one can write
\[ dx^a = \frac{\partial x^a}{\partial x^1} dx^1 + \frac{\partial x^a}{\partial x^2} dx^2 \cdots \frac{\partial x^a}{\partial x^N} dx^N = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} dx^b, \]  
(17.33)
where the summation is evaluated at \( P \). Similarly for the interval between \( x^a \) and \( x^a + dx^a \), in the none-primed coordinate system, we can write
\[ dx^a = \frac{\partial x^a}{\partial x^1} dx^1 + \frac{\partial x^a}{\partial x^2} dx^2 \cdots \frac{\partial x^a}{\partial x^N} dx^N = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} dx^b, \]  
(17.34)
here also the summation is evaluated at \( P \).

*Einstein’s summation convention:* whenever an index occurs twice in an expression, once as subscript and once as a superscript, imply a summation over the index. An index should not occur more than twice. For example, according to Einstein’s summation convention, the summations in Eq. (17.33) and (17.34) can be expressed
\[ dx^a = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} dx^b = g_{b}^{a}(x) dx^b = \frac{\partial x^a}{\partial x^b} dx^b \]  
(17.35)
and
\[ dx^a = \sum_{b=1}^{N} \frac{\partial x^a}{\partial x^b} dx^b = g_{b}^{a}(x') dx^b = \frac{\partial x^a}{\partial x^b} dx^b. \]  
(17.36)
The index $a$, which is known as the *free index*, can take any value from 1 to $N$. The index $b$, which is known as the *dummy index* and it must be summed up from 1 to $N$.

### 17.4 The Riemannian geometry

*The local geometry of a Manifold:* The local geometry of a manifold is determined by defining the invariant 'distance' or (as we saw in the Minkowski spacetime Manifold) the interval $ds$ between points $P$ with coordinate $x^a$ and $Q$ with coordinates $x^a + dx^a$. This distance can be assigned in general to be a well-behaved function $g(x^a, dx^a)$ of the coordinates $x^a$ and $dx^a$.

$$ds^2 = g(x^a, dx^a).$$  \hfill (17.37)

Let’s reconsider the worldline of the alien in the spaceship in the previous chapter. Imagine there is a surface defined by the two coordinates $(x, ct)$ represented as $(x^1, x^2)$. Let’s consider two points $P$ and $Q$, as shown in Fig. 17.7. The coordinates for the points are $(x^1, x^2)$ and $(x^1 + dx^1, x^2 + dx^2)$, respectively. Suppose this surface is defined by the function $s(x^1, x^2)$ The interval between these two points, $ds$, can be expressed as

$$ds = \frac{\partial s}{\partial x^1} dx^1 + \frac{\partial s}{\partial x^2} dx^2$$

$$\Rightarrow ds^2 = \left(\frac{\partial s}{\partial x^1} dx^1 + \frac{\partial s}{\partial x^2} dx^2\right) \left(\frac{\partial s}{\partial x^1} dx^1 + \frac{\partial s}{\partial x^2} dx^2\right)$$

$$\Rightarrow ds^2 = \frac{\partial s}{\partial x^1} \frac{\partial s}{\partial x^1} dx^1 dx^1 + \frac{\partial s}{\partial x^1} \frac{\partial s}{\partial x^2} dx^1 dx^2 + \frac{\partial s}{\partial x^2} \frac{\partial s}{\partial x^1} dx^1 dx^2$$

$$+ \frac{\partial s}{\partial x^2} \frac{\partial s}{\partial x^2} dx^2 dx^2$$  \hfill (17.38)
17.4. THE RIEMANNIAN GEOMETRY

This can be rewritten as

\[
    ds^2 = g_{11}(x^1, x^2) \, dx^1 \, dx^1 + g_{12}(x^1, x^2) \, dx^1 \, dx^2 + g_{21}(x^1, x^2) \, dx^1 \, dx^3 + g_{22}(x^1, x^2) \, dx^2 \, dx^2 = 2 \sum_{a=1}^{2} \sum_{b=1}^{2} g_{ab}(x) \, dx^a \, dx^b \tag{17.39}
\]

where we replaced

\[
    g_{11} = \frac{\partial s^2}{\partial x^1 \partial x^1}, \quad g_{12} = \frac{\partial s^2}{\partial x^1 \partial x^2}, \quad g_{21} = \frac{\partial s^2}{\partial x^2 \partial x^1}, \quad g_{22} = \frac{\partial s^2}{\partial x^2 \partial x^2}
\]

Consider a curved surface in 3-D Euclidean space. We know that this surface can be defined by a function \( s(x) \), that depends on \((x, y, z)\) in Cartesian, \((r, \theta, \varphi)\) in spherical, or \((r, \varphi, z)\) in cylindrical coordinates. Suppose we represent these coordinates by \(x_1, x_2, x_3\), then we may write the function that defines the surface as

\[
    s(x) = g(x) = g(x^1, x^2, x^3). \tag{17.40}
\]

Now let’s consider a point \(P\) on this surface that has coordinates \((x_1, x_2, x_3)\). Suppose we consider another point \(Q\) with coordinates \((x_1 + dx_1, x_2 + dx_2, x_3 + dx_3)\), we may define the surface between these two points as \(ds^2\). This displacement is just the differential of Eq. (17.40)

\[
    ds^2 = \left( \frac{\partial g}{\partial x^1} \, dx^1 + \frac{\partial g}{\partial x^2} \, dx^2 + \frac{\partial g}{\partial x^3} \, dx^3 \right) \cdot \left( \frac{\partial g}{\partial x^1} \, dx^1 + \frac{\partial g}{\partial x^2} \, dx^2 + \frac{\partial g}{\partial x^3} \, dx^3 \right)
\]

or using Einstein’s summation convention, the geometry of the surface between the two points

\[
    ds^2 = (ds) \cdot (ds) = \left( \sum_{a=1}^{3} \frac{\partial g}{\partial x^a} \, dx^a \right) \cdot \left( \sum_{b=1}^{3} \frac{\partial g}{\partial x^b} \, dx^b \right)
\]

\[
    = \sum_{a=1}^{3} \sum_{b=1}^{3} \frac{\partial g}{\partial x^a} \frac{\partial g}{\partial x^b} \, dx^a \, dx^b. \tag{17.42}
\]

Again using Einstein’s summation convention and the notation

\[
    g_{ab}(x) = \frac{\partial g}{\partial x^a} \frac{\partial g}{\partial x^b}
\]

we may write

\[
    ds^2 = g_{ab}(x) \, dx^a \, dx^b. \tag{17.44}
\]

which is the metric equation that we defined earlier. Note that the metric tensor in this case is a \(3 \times 3\) matrix given by

\[
    G = \begin{bmatrix}
        \frac{\partial g}{\partial x^1} & \frac{\partial g}{\partial x^2} & \frac{\partial g}{\partial x^3} \\
        \frac{\partial g}{\partial x^2} & \frac{\partial g}{\partial x^1} & \frac{\partial g}{\partial x^3} \\
        \frac{\partial g}{\partial x^3} & \frac{\partial g}{\partial x^3} & \frac{\partial g}{\partial x^3}
    \end{bmatrix} \tag{17.45}
\]
and we can easily see that this matrix is symmetric as

$$\frac{\partial g}{\partial x^a} \frac{\partial g}{\partial x^b} = \frac{\partial g}{\partial x^b} \frac{\partial g}{\partial x^a}.$$ 

Suppose the coordinates for an orthonormal set like the Cartesian, spherical, or cylindrical, we note that

$$G = \begin{pmatrix} \left(\frac{\partial g}{\partial x^1}\right)^2 & 0 & 0 \\ 0 & \left(\frac{\partial g}{\partial x^2}\right)^2 & 0 \\ 0 & 0 & \left(\frac{\partial g}{\partial x^3}\right)^2 \end{pmatrix}$$

(17.46)

In general theory of relativity, we are interested in a Manifold where the interval $ds$ can be described by the equation of the form

$$ds^2 = \sum_{a=1}^{N} \sum_{b=1}^{N} g_{ab}(x) dx^a dx^b,$$

(17.47)

or simply using the Einstein’s summation convention

$$ds^2 = g_{ab}(x) dx^a dx^b.$$ 

(17.48)

A geometry of a Manifold defined by Eq. (17.48) is known as the Riemannian geometry if $ds^2 > 0$. As we have seen in the case of Minkowski spacetime manifold the interval $ds^2$ can also be negative (spacelike) or zero (lightlike). In such cases the geometry is referred as pseudo Riemannian geometry and the manifold can be referred as pseudo Riemannian. The function $g(x)$ is known as the metric function, where $g_{ab}(x)$ represent the element of a metric tensor.

**Transformation of the interval:** applying the relation in Eq. (17.36), we can express

$$dx^a = \frac{\partial x^a}{\partial x'^c} dx'^c, \quad dx^b = \frac{\partial x^b}{\partial x'^d} dx'^d,$$ 

(17.49)

so that the interval can be transformed as

$$ds^2 = g_{ab}(x) \frac{\partial x^a}{\partial x'^c} \frac{\partial x^b}{\partial x'^d} dx'^c dx'^d,$$ 

(17.50)

or

$$ds^2 = g'_{cd}(x') dx'^c dx'^d,$$ 

(17.51)

where

$$g'_{cd}(x') = g_{ab}(x) \frac{\partial x^a}{\partial x'^c} \frac{\partial x^b}{\partial x'^d}.$$ 

(17.52)

Note that $x = x(x').$
17.5 Intrinsic and extrinsic geometry and the metric

A given geometry of dimension $M$ defined by the metric equation

$$ds^2 = g_{ab}(x) dx^a dx^b.$$  \hspace{1cm} (17.53)

and embedded in a higher dimension manifold of dimension $N$, ($N > M$) is said to be:

(a) Intrinsic: when the geometry remains unchanged as viewed in the higher dimensional manifold.

(b) Extrinsic: when the geometry is different as viewed in the higher dimensional manifold.

**Example 2.2 Extrinsic geometry:** One simple example of extrinsic geometry is 2-D cylindrical geometry. In order to see that let’s consider a plane geometry in a 3-D Euclidean Manifold. Let’s this plane depends on $(x^1, x^2)$. We recall from theoretical physics I generally a plane:

*Equation of a plane:* If $\vec{N} = ai + bj + ck$ is normal (perpendicular) to a plane, then the scalar product of the vector $\vec{N}$ and the vector $\vec{r} - \vec{r}_0$

$$\vec{r} - \vec{r}_0 = (x - x_0) \hat{x} + (y - y_0) \hat{y} + (z - z_0) \hat{z}$$ \hspace{1cm} (17.54)

is zero,

$$\vec{N} \cdot (\vec{r} - \vec{r}_0) = a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$ \hspace{1cm} (17.55)

This defines the equation of the plane. It can be rewritten as

$$ax + by + cz = d.$$ \hspace{1cm} (17.56)
where
\[ d = ax_0 + by + cz_0. \]  
(17.57)

Using the notation \((x^1, x^2, x^3)\) for \((x, y, z)\), we may write
\[ ax^1 + bx^2 + cx^3 = d, \]  
(17.58)

where
\[ d = ax^1_0 + bx^2_0 + cx^3_0. \]  
(17.59)

For a plane that depends on only \((x^1, x^2)\) we may write the interval between
\( (x^1, x^2) \) and \((x^1 + dx^1, x^2 + dx^2) \) as
\[ ds^2 = (dx^1)^2 + (dx^2)^2. \]  
(17.60)

Now let’s consider the 2-D cylindrical surface which we can construct using our 2-D plane. Suppose the cylinder has radius, \(a\), with its axis along the \(z\)-axis (which we call \(z^3\) - axis). Using cylindrical coordinates a point on the surface of the cylinder can be described by \((a, \varphi, z)\) or using our coordinates notations \((a, x^2, x^3)\), we can define the surface by the function
\[ g(a, \varphi, z) = g(a, x^2, x^3). \]  
(17.61)

we may write the interval between two points on this surface, point \(P\) and \(Q\) with coordinates \((x^1, x^2, x^3)\) and point \(Q\) with coordinates \((x^1 + dx^1, x^2 + dx^2, x^3 + dx^3)\) just using geometrical visualization, as
\[ ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2. \]  
(17.62)

For the cylinder with radius \(a\) shown in Fig. 17.9
\[ x^1 = a \cos (x^2), \]  
\[ x^2 = a \sin (x^2), \]  
\[ x^3 = x^3 \]
we find
\[ ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = a^2 (dx^2)^2 + (dx^3)^2. \]  
(17.63)
17.5. INTRINSIC AND EXTRINSIC GEOMETRY AND THE METRIC 499

This is a 2-D surface embedded in a 3-D manifold. It has cylindrically curved geometry when it is viewed in this 3-D Euclidean manifold. But you can actually obtain this geometry from the plane geometry from Eq. (17.40) by simply substituting

\[ x^1 = a x^2, x^2 = x'^3 \]

Such kind of geometry is not intrinsic and it is called extrinsic. Its curvature is extrinsic and is a result of the way it is embedded in the three dimensional space..

**Example 2.3 Intrinsic geometry:** One simple example of intrinsic geometry is 2-D spherical geometry embedded in a 3-D Euclidean Manifold. In Fig. 17.5 we see a 3-D infinitesimal volume. Assume a sphere with radius \( a \).

Then the surface defined by a pair of points on this sphere separated by a distance \( ds \) can be expressed as

\[ ds^2 = (ad\theta)^2 + (a \sin(\theta) d\phi)^2 = a^2 (d\theta)^2 + a^2 \sin^2(\theta) (d\phi)^2. \quad (17.64) \]
or using the notation \((x^2, x^3)\) for \((\theta, \varphi)\), we can write

\[
ds^2 = a^2 (dx^2)^2 + a^2 \sin^2 (x^2) (dx^3)^2. \tag{17.65}
\]

This is a 2-D surface embedded in a 3-D manifold. You can not obtain this geometry from the plane geometry like the 2-D cylindrical geometry. Such kind of geometry is *intrinsic*. This means the geometry of a sphere is intrinsically curved because we can not transform Eq.(17.65) to the Euclidean form

\[
ds^2 = (dx^1)^2 + (dx^2)^2 \tag{17.66}
\]

over the whole surface by any coordinate transformation. Note that this can be done locally but not for the whole spherical surface.

**Example 2.4** Find the metric for a two-dimensional sphere of radius, \(a\), embedded in a 3-D Euclidean space both in Cartesian coordinates \((x^1, x^2, x^3)\). Refer to Fig. 17.5

\[
\begin{align*}
\text{Solution:} & \quad \text{We recall that the line element in a 3-D Euclidean space is given by} \\
& \quad ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2, \tag{17.67}
\end{align*}
\]

For a two dimensional spherical geometry with radius \(a\), we have

\[
(x^1)^2 + (x^2)^2 + (x^3)^2 = a^2, \tag{17.68}
\]

Using (17.68), we can write

\[
x^3 = \sqrt{a^2 - (x^1)^2 - (x^2)^2}, \tag{17.69}
\]
so that
\[
dx^3 = -\frac{x^1 dx^1 + x^2 dx^2}{\sqrt{a^2 - (x^1)^2 + (x^2)^2}}.
\]

Then for a 2-D sphere embedded in a 3-D Euclidean space, the metric is given by
\[
ds^2 = (dx^1)^2 + (dx^2)^2 + \left(\frac{x^1 dx^1 + x^2 dx^2}{\sqrt{a^2 - (x^1)^2 + (x^2)^2}}\right)^2.
\]

If we consider a point in the neighborhood of the pole, we may set \(x^1 = x^2 \approx 0\) and the metric in Eq. (17.70) reduces to the Euclidean form
\[
ds^2 = (dx^1)^2 + (dx^2)^2.
\]

Let’s use the coordinates \((x'^1, x'^2, x'^3)\), defined by the transformation
\[
x^1 = x'^1 \cos (x'^2), \quad x^2 = x'^1 \sin (x'^2), \quad x^3 = x'^3
\]

For any point on the surface of the sphere, we have
\[
(x^1)^2 + (x^2)^2 + (x^3)^2 = a^2
\]
\[
\Rightarrow x^3 = \sqrt{a^2 - (x'^1)^2 + (x'^2)^2} = \sqrt{a^2 - (x'^1)^2},
\]

Note that the origin is set at the north pole of the sphere at point as shown in Fig. 17.5.,. Then the interval can be written as
\[
ds^2 = (dx'^1 \cos (x'^2) - x'^1 \sin (x'^2) dx'^2)^2
\]
\[
+ (dx'^1 \sin (x'^2) + x'^1 \cos (x'^2) dx'^2)^2 + \left(\frac{x'^1 dx'^1}{\sqrt{a^2 - (x'^1)^2}}\right)^2
\]
so that after a little algebra, we find
\[
ds^2 = (dx'^1)^2 + (x'^1)^2 \left(dx'^2\right)^2 + \left(\frac{x'^1 dx'^1}{\sqrt{a^2 - (x'^1)^2}}\right)^2
\]

which simplifies into
\[
ds^2 = \frac{a^2 (dx'^1)^2}{a^2 - (x'^1)^2} + (x'^1)^2 \left(dx'^2\right)^2
\]
\[ ds^2 = g_{11} (dx^1)^2 + g_{22} (dx^2)^2 \]  

(17.76)

where

\[ g_{11} = \frac{a^2}{a^2 - (x^1)^2}, g_{22} = (x^1)^2 \]  

(17.77)

are the none zero elements of the metric tensor. We will see the use of these elements of the metric tensor in the next section to determine length and area of a 2-D sphere in a 3-D Euclidean manifold.

**Example 2.5** Determine the metric for a three-dimensional sphere of radius \( a \) embedded in a 4-D Euclidean space: We can write the interval

\[ ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + (dx^4)^2, \]  

(17.78)

For a three dimensional spherical geometry with radius \( a \) we have

\[ (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 = a^2 \]

\[ \Rightarrow x^4 = \sqrt{a^2 - (x^1)^2 - (x^2)^2 - (x^3)^2} \]  

(17.79)

so that

\[ dx^4 = -\frac{x^1 dx^1 + x^2 dx^2 + x^3 dx^3}{\sqrt{a^2 - (x^1)^2 - (x^2)^2 - (x^3)^2}}. \]

Then the metric can be expressed as

\[ ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + \frac{(x^1 dx^1 + x^2 dx^2 + x^3 dx^3)^2}{a^2 - (x^1)^2 - (x^2)^2 - (x^3)^2}, \]  

(17.80)

We introduce the coordinates \((r, \theta, \varphi)\) which we represent \((x^1, x^2, x^3)\).

and defined by the transformation

\[ x^1 = r \sin(\theta) \cos(\varphi), x^2 = r \sin(\theta) \sin(\varphi), x^3 = \cos(\theta), \]

or

\[ x^1 = x'^1 \sin(x'^3), x^2 = \sin(x'^2) \sin(x'^3), x^3 = \cos(x'^2). \]

Now referring to Fig.17.5 the "distance" squared between point \( P \) and \( Q \) can easily be determined using the Pythagorean theorem. First find the length of the hypotenuse of the green triangle. Suppose if we call this length \( ds_s \), we note that

\[ ds_s^2 = (r \sin(\theta) d\varphi)^2 + (rd\theta)^2. \]  

(17.81)
Then the distance between $P$ and $Q$ shown by the red line (hypotenuse side) can be expressed as

$$ds^2 = dr^2 + ds_s^2 = dr^2 + r^2 \sin^2(\theta) d\varphi^2 + r^2 d\theta^2$$  \hspace{1cm} (17.82)

Note that I referred this distance as $ds'$ because it represents distance

$$ds'^2 = dx^2 + dy^2 + dz^2$$

We are considering a 3-D sphere in a 4-D Euclidean manifold, where the interval between point $P$ and $Q$ is given by

$$ds^2 = ds'^2 + (dx^4)^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + (dx^4)^2,$$ \hspace{1cm} (17.83)

with the constraint

$$x^4 = \sqrt{a^2 - (x^1)^2 - (x^2)^2 - (x^3)^2}.$$ \hspace{1cm} (17.84)

which we may write, in terms of the polar coordinates, as

$$x^4 = \sqrt{a^2 - r^2}.$$ \hspace{1cm} (17.84)

so that

$$dx^4 = -\frac{r dr}{\sqrt{a^2 - r^2}}.$$ \hspace{1cm} (17.85)

Then the metric for a 3-D sphere in a 4-D Euclidean manifold is given by

$$ds^2 = ds'^2 + dx^4 = dr^2 + r^2 \sin^2(\theta) d\varphi^2 + r^2 d\theta^2 + \frac{r^2 dr^2}{a^2 - r^2},$$ \hspace{1cm} (17.86)
which can be rewritten as

\[ ds^2 = \frac{a^2}{a^2 - r^2} dr^2 + r^2 d\theta^2 + r^2 \sin^2(\theta) d\varphi^2 \]  \hspace{1cm} (17.87)

or

\[ ds^2 = g_{11} \,(dx^1)^2 + g_{22} \,(dx^2)^2 + g_{33} \,(dx^3)^2 , \]  \hspace{1cm} (17.88)

where we used the notation \((x^1, x^2, x^3)\) for the coordinates \((r, \theta, \varphi)\) and identify the none zero elements of the metric tensor

\[ g_{11} = \frac{a^2}{a^2 - r^2}, \quad g_{22} = r^2, \quad g_{33} = r^2 \sin^2(\theta) \]

We will see the use of these elements of the metric tensor in the next section to determine length, area, and volume of a 3-D sphere in a 4-D Euclidean manifold.

### 17.6 Length, area, and volume

**Length:** Suppose two points \(P\) and \(Q\) on a manifold of dimension \(N\) are connected by some curve. The length of the curve connecting these two points is given by

\[ L_{PQ} = \int_P^Q \sqrt{|ds^2|} = \int_P^Q \sqrt{\left|g_{ab}(x)\,dx^a dx^b\right|} \]  \hspace{1cm} (17.89)

the absolute value is because of the fact that for pseudo-Riemannian manifolds it can be negative as is the case for spacelike in the Minkawski spacetime manifold. For \(x^a = x^a\,(u)\), we recall write

\[ dx^a = \frac{dx^a}{du} du, \quad dx^b = \frac{dx^b}{du} du \]  \hspace{1cm} (17.90)

so that

\[ L_{PQ} = \int_P^Q \sqrt{\left|g_{ab}(u)\,\frac{dx^a}{du}\frac{dx^b}{du}\right|} du. \]  \hspace{1cm} (17.91)

**Area and volume:**

Generally the area can determined using

\[
A = \iiint ds_1 \times ds_2 = \iiint \sqrt{|g_{ab}(x)\,dx^a dx^b|} \times \sqrt{|g_{cd}(x)\,dx^c dx^d|}, \text{ for } c, d \neq a, b
\]

\[
= \iiint \sqrt{|g_{ab}(x)\,g_{cd}(x)\,dx^a dx^b dx^c dx^d|} = \iiint \sqrt{|g_{ab}(x)\,g_{cd}(x)\,\frac{dx^a}{dx^b}\,\frac{dx^c}{dx^d}\,(dx^b dx^d)^2|}
\]

\[
\Rightarrow A = \iiint \sqrt{|g_{ab}(x)\,g_{cd}(x)\,\frac{dx^a}{dx^b}\,\frac{dx^c}{dx^d}|} \,dx^b dx^d. \]  \hspace{1cm} (17.92)
For an orthogonal set of coordinates, \( x^p \)
\[
\frac{dx^p}{dx^q} = \delta^{pq} = \begin{cases} 
1, & p = q \\
0, & p \neq q
\end{cases}
\] (17.93)
and the area becomes
\[
A = \iiint \sqrt{g_{ab}(x) g_{cd}(x) \delta^{ac} \delta^{bd}} \, dx^b \, dx^d = \iiint \sqrt{|g_{bb}(x) g_{dd}(x)|} \, dx^b \, dx^d ,
\]
note that, \( d \neq b \). As an example let’s consider a 2D hypersurface embedded in a 3D manifold. Then the surface area is given by
\[
A = \sum_{i=1}^{2} \left\{ \iiint \sqrt{|g_{11}(x) g_{dd}(x)|} \, dx^1 \, dx^d + \iiint \sqrt{|g_{22}(x) g_{dd}(x)|} \, dx^2 \, dx^d \right. \\
+ \iiint \sqrt{|g_{33}(x) g_{dd}(x)|} \, dx^3 \, dx^d \right\}.
\]
\[
= \iiint \sqrt{|g_{22}(x) g_{11}(x)|} \, dx^2 \, dx^1 + \iiint \sqrt{|g_{33}(x) g_{11}(x)|} \, dx^3 \, dx^1 \\
+ \iiint \sqrt{|g_{11}(x) g_{22}(x)|} \, dx^1 \, dx^2 + \iiint \sqrt{|g_{33}(x) g_{22}(x)|} \, dx^3 \, dx^d \\
+ \iiint \sqrt{|g_{11}(x) g_{33}(x)|} \, dx^1 \, dx^3 + \iiint \sqrt{|g_{22}(x) g_{33}(x)|} \, dx^2 \, dx^3
\]
For a hypersurface, like the 2D sphere in the 3D manifold, the surface is defined by the equation
\[
x^3 = \text{constant} \Rightarrow dx^3 = 0.
\]
and the expression for the area reduces to
\[
A = \iiint \sqrt{|g_{22}(x) g_{11}(x)|} \, dx^2 \, dx^1 + \iiint \sqrt{|g_{11}(x) g_{22}(x)|} \, dx^1 \, dx^2
\]
which can simply be written as
\[
A = \iiint \sqrt{|g_{22}(x) g_{11}(x)|} \, dx^2 \, dx^1
\]
where one can absorb the factor 2 into the metric elements. For such coordinates the metric is diagonal
\[
ds^2 = g_{11}(dx^1)^2 + g_{22}(dx^2)^2 + \ldots + g_{NN}(dx^N)^2.
\] (17.94)
and the infinitesimal area \( dA \) of the surface defined by
\[
x^a = \text{constant},
\] (17.95)
for, \( a = 3, 4 \ldots N \), is given by
\[ dA = \sqrt{|g_{11}g_{22}|} dx^1 dx^2 \]  
(17.96)

and for 3-D volume in the \((x^1, x^2, x^3)\) defined by

\[ x^a = \text{constant} \]  
(17.97)

for \(a = 4, 5, \ldots, N\), the infinitessimal volume is given by

\[ dV = \sqrt{|g_{11}g_{22}g_{33}|} dx^1 dx^2 dx^3 \]  
(17.98)

**Example 2.5** For the two-dimensional sphere of radius \(a\) embedded in a 3-D Euclidean space consider a surface defined by a radius, \(\rho = R\). For this surface find

(a) The distance, \(D\), from the origin to the perimeter on this surface along a line of constant \(\varphi\) \(i.e.\ x^2 = \text{cons}\) The origin is at the north pole as shown in Fig. 17.6.

(b) The circumference of the circle with radius, \(\rho = R\) \(i.e.\ x^1 = R = \text{cons}\).

(c) The area of the spherical surface enclosed by the perimeter with radius, \(\rho = R\), \(i.e.\ the surface shaded green in Fig. 17.6)\.

**Solution:**

(a) We recall that the metric for a 2-D sphere in a 3-D Euclidean manifold is given by

\[ ds^2 = g_{11} (dx^1)^2 + g_{22} (dx^2)^2, \]  
(17.99)

where we used the notation \((x^1, x^2)\) for the coordinates \((\rho, \varphi)\) and identify the none zero elements of the metric tensor

\[ g_{11} = \frac{a^2}{a^2 - \rho^2}, \quad g_{22} = \rho^2. \]  
(17.100)
We recall the length of a curve between two points, $P$ and $Q$, on a manifold in terms of the metric tensor is given by

$$
L_{PQ} = \int_{P}^{Q} \sqrt{|g_{ab}(x) dx^a dx^b|}. \quad (17.101)
$$

For curve on the surface of the sphere shown in red in Fig. 17.6, the coordinates for point $P$ is $(x^1 = \rho = 0, x^2 = \phi = \text{constant})$ and for point $Q (x^1 = \rho = R, x^2 = \phi = \text{constant})$ which leads to $dx^2 = d\phi = 0$.

Then the length becomes

$$
D = \int_{P}^{Q} \sqrt{|g_{11}(x^1) dx^1 dx^1|} = \int_{0}^{R} \frac{a}{\sqrt{a^2 - \rho^2}} d\rho = a \sin^{-1} \left[ \frac{R}{a} \right]. \quad (17.102)
$$

(b) Along the circumference (the curve shown in pink Fig. 17.6), we know that $x^1 = \rho = R = \text{constant}$ which leads to $dx^1 = d\rho = 0$.

Thus the expression

$$
L_{PQ} = \int_{P}^{Q} \sqrt{|g_{ab}(x^1) dx^a dx^b|} \quad (17.103)
$$

for the circumference becomes

$$
C = \int_{0}^{2\pi} \sqrt{|g_{22}(x^1) (dx^2)^2|} = \int_{0}^{2\pi} \sqrt{|g_{22}(x^1 = \rho = R)| (dx^2)}
$$

$$
\Rightarrow C = \int_{0}^{2\pi} R dx^2 = 2\pi R \quad (17.104)
$$
(c) The surface is defined by $x^3 =$ constant. Then according to Eq. (17.96), the infinitesimal area is given by
\[
dA = \sqrt{|g_{11}g_{22}|} dx^1 dx^2.
\]
Then the surface area becomes
\[
A = \int\int \sqrt{\frac{a^2}{a^2 - \rho^2}} d\rho d\phi = \int\int \frac{a}{\sqrt{a^2 - \rho^2}} d\rho d\phi.
\]
(17.107)

To cover the area (colored Aqua) shown in Fig 17.6, we should have for the limits of integrations $[0, R]$ for $\rho$ and $[0, 2\pi]$ for $\varphi$
\[
A = \int_0^R \int_0^{2\pi} \frac{a}{\sqrt{a^2 - \rho^2}} d\rho d\phi = 2\pi a^2 \left[ 1 - \sqrt{1 - \frac{R^2}{a^2}} \right].
\]
(17.108)

_Homework:_ In Example 2.5, express the circumference (part b) and the area (part c) in terms of the distance $D$ (part a) and determine $D$ for the maximum circumference and area. Find the maximum circumference and area of this sphere. Is your answer is consistent with what you know about the maximum circumference and surface area of a sphere with radius $a$.

**Example 2.6** For the three-dimensional sphere embedded in a 4-D Euclidean space by considering a 2-D sphere of coordinate radius $r = R$, find

(a) the distance from its center to the surface of this sphere along constant $\theta$ and constant $\phi$.

(b) the circumference across the equator
(c) the area of the spherical surface.

(d) The volume bounded by the spherical surface.

Solution:

(a) We recall the metric for a 3-D sphere embedded in a 4-D Euclidean space is given by

\[ ds^2 = g_{11}(dx^1)^2 + g_{22}(dx^2)^2 + g_{33}(dx^3)^2, \]  

(17.109)

where we used the notation \((x^1, x^2, x^3)\) for the coordinates \((r, \theta, \varphi)\) and identify the none zero elements of the metric tensor

\[ g_{11} = \frac{a^2}{a^2 - r^2}, g_{22} = r^2, g_{33} = r^2 \sin^2(\theta) \]  

(17.110)

or in terms of \((x'^1, x'^2, x'^3)\)

\[ g_{11} = \frac{a^2}{a^2 - (x'^1)^2}, g_{22} = (x'^1)^2, g_{33} = (x'^1)^2 \sin^2(\theta) . \]  

(17.111)

For \(x'^2 = \theta = \text{constant}\) and \(x'^3 = \phi = \text{constant}\), we have

\[ dx'^2 = dx'^3 = 0. \]  

(17.112)

so that the distance \(D\)

\[ L_{PQ} = \int_P^Q \sqrt{|g_{ab}(x') dx'^a dx'^b|} \]  

(17.113)

becomes

\[ L_{PQ} = \int_P^Q \sqrt{|g_{11}(x') dx'^1 dx'^1|} \]  

(17.114)

or

\[ D = \int_0^R \sqrt{ \frac{a^2}{a^2 - r^2} dr^2 } = \int_0^R \frac{a}{\sqrt{a^2 - r^2}} dr = a \sin^{-1} \left[ \frac{R}{a} \right] . \]  

(17.115)

which is the same as the result we obtained in the previous example.

(b) Across the equator, we have \(x'^1 = r = R\), \(x'^2 = \theta = \pi/2\) and obviously \(dx'^1 = 0\) and \(dx'^2 = 0\). Then the length

\[ L_{PQ} = \int_P^Q \sqrt{|g_{ab}(x') dx'^a dx'^b|}, \]  

(17.116)

for the circumference, becomes

\begin{align*}
C &= \int \sqrt{|g_{33}(x') (dx'^3)^2|} = \int_0^{2\pi} \sqrt{|g_{33}(x'^1 = r = R, x'^2 = \pi/2)|} dx'^3 \\
&\Rightarrow C = \int_0^{2\pi} \sqrt{R^2 \sin^2(\pi/2)} dx'^3 = \int_0^{2\pi} R dx'^3 = 2\pi R \end{align*}  

(17.117)
(c) The spherical surface is defined by \( x'^1 = R = \text{constant} \). Therefore, in view of Eq. (17.96), the infinitesimal area on this surface should be expressed as
\[
dA = \sqrt{|g_{22}g_{33}|} dx'^2 dx'^3.
\] (17.118)
Noting that for \( x'^1 = r = R \)
\[
g_{22} = r^2 \bigg|_{r=R} = R^2
\]
\[
g_{33} = r^2 \sin^2(\theta) \bigg|_{r=R} = R^2 \sin^2(\theta)
\] (17.119)
and the limit of integration for \( x'^2 \) is \((0, \pi)\) and for \( x'^3 \) is \((0, 2\pi)\), the surface area would be become
\[
A = \int_0^\pi \int_0^{2\pi} \sqrt{R^4 \sin^2(\theta)} \, d\theta d\phi = \int_0^\pi \int_0^{2\pi} R^2 \sin(\theta) \, d\theta d\phi = 4\pi R^2.
\] (17.120)

(d) In this case we are considering a 3-D sphere embedded in a 4-D Euclidean space. For this space the volume is defined by \( x'^4 = \text{constant} \). Therefore, applying the relation in Eq. (17.98), an infinitesimal volume in this 4-D space is given by
\[
dV = \sqrt{|g_{11}g_{22}g_{33}|} dx'^1 dx'^2 dx'^3.
\] (17.121)
Thus using
\[
g_{11} = \frac{a^2}{a^2 - r^2}, g_{22} = r^2, g_{33} = r^2 \sin^2(\theta)
\] (17.122)
the volume bounded by the 2-D spherical surface of radius \( x'^1 = r = R \) becomes
\[
V = \int_0^R \int_0^\pi \int_0^{2\pi} \sqrt{\frac{a^2 r^4 \sin^2(\theta)}{a^2 - r^2}} \, dr d\theta d\phi
\]
\[
= a \int_0^R \frac{r^2 \, dr}{\sqrt{a^2 - r^2}} \int_0^\pi \int_0^{2\pi} \sin(\theta) \, d\theta d\phi
\] (17.123)
or
\[
V = a \int_0^R \frac{(x'^1)^2 \, dx'^1}{\sqrt{a^2 - (x'^1)^2}} \int_0^\pi \int_0^{2\pi} \frac{\sin(x'^2) \, dx'^2 dx'^3}{\sqrt{a^2 - (x'^1)^2}}.
\]
Let’s evaluate the integral
\[
I = \int \frac{r^2 \, dr}{\sqrt{a^2 - r^2}}
\] (17.124)
Introducing the transformation defined by
\[
r = a \sin(\theta) \Rightarrow dr = a \cos(\theta) \, d\theta
\]
we have
\[ I = \int \frac{r^2 dr}{\sqrt{a^2 - r^2}} = \int \frac{a^2 \sin^2 (\theta) a \cos (\theta) d\theta}{\sqrt{a^2 - a^2 \sin^2 (\theta)}} = a^2 \int \sin^2 (\theta) d\theta \]
\[ = \frac{a^2}{2} \int (1 - \cos (2\theta)) d\theta = \frac{a^2}{2} \left( \theta - \frac{\sin (2\theta)}{2} \right) \]
\[ \Rightarrow I = \int \frac{r^2 dr}{\sqrt{a^2 - r^2}} = \frac{a^2}{2} \left[ \theta - \sin (\theta) \cos (\theta) \right] \quad (17.125) \]
so that using
\[ r = a \sin (\theta) \Rightarrow \sin (\theta) = \begin{cases} \frac{R}{a}, & \text{for } r = R \\ 0, & \text{for } r = 0 \end{cases} \quad (17.126) \]
and
\[ \cos (\theta) = \sqrt{1 - \sin^2 (\theta)} = \begin{cases} \sqrt{1 - \left( \frac{R}{a} \right)^2}, & \text{for } r = R \\ 0, & \text{for } r = 0 \end{cases} \quad (17.127) \]
one finds
\[ \int_0^R \frac{r^2 dr}{\sqrt{a^2 - r^2}} = \frac{a^2}{2} \left[ \sin^{-1} \left( \frac{R}{a} \right) - \frac{R}{a} \sqrt{1 - \left( \frac{R}{a} \right)^2} \right] \quad (17.128) \]
Then the volume becomes
\[ V = 2\pi a^3 \left\{ \sin^{-1} \left( \frac{R}{a} \right) - \frac{R}{a} \sqrt{1 - \left( \frac{R}{a} \right)^2} \right\} \quad (17.129) \]
One must be able to recover the 3-D Euclidean space for \( a \to \infty \). This means \( \frac{R}{a} \ll 1 \) the result for the volume of a sphere with radius, \( R \) must be that of the volume of a sphere with radius \( R \) in 3D Euclidean space \((V = \frac{4}{3}\pi R^3)\). One can easily find from Eq. (17.129) using the approximations for \( \frac{R}{a} \ll 1 \),
\[ \sin^{-1} \left( \frac{R}{a} \right) \simeq \frac{R}{a} - \frac{1}{3} \left( \frac{R}{a} \right)^3, \sqrt{1 - \left( \frac{R}{a} \right)^2} = 1 - \frac{1}{2} \left( \frac{R}{a} \right)^2 \quad (17.130) \]
that gives
\[ V = 2\pi a^3 \left\{ \frac{R}{a} - \frac{1}{3} \left( \frac{R}{a} \right)^3 - \frac{R}{a} \left( 1 - \frac{1}{2} \left( \frac{R}{a} \right)^2 \right) \right\} \]
\[ = 2\pi a^3 \left\{ -\frac{1}{3} \left( \frac{R}{a} \right)^3 + \frac{1}{2} \left( \frac{R}{a} \right)^3 \right\} \Rightarrow V = \frac{4}{3}\pi R^3 \quad (17.131) \]
CHAPTER 17. MANIFOLDS

Homework:

(a) Express the circumference, the area, and the volume in terms of $D$ in the previous example and show that all have maximum values at

$$D = \frac{\pi a}{2}$$  \hspace{1cm} (17.132)

(b) Show that the total volume of this space is finite and is equal to

$$V = 2\pi^2 a^3.$$  \hspace{1cm} (17.133)

Homework: Determine the metric for a three-dimensional sphere with imaginary radius $a = ib$ embedded in a 4-D Euclidean space: and by considering a sphere defined by $r = R$ find

(a) Show that circumference, $C$, and the area, $A$ are still $C = 2\pi R$ and $A = 4\pi R^2$.

(b) The distance $D$ from the center of the sphere to the surface is

$$D = b \sinh^{-1} \left(\frac{R}{b}\right)$$  \hspace{1cm} (17.134)

and in this case show that $A$ and $V$ of the sphere are monotonically increasing functions

17.7 Local Cartesian coordinates and tangent space

Generally $ds^2$, can be positive, negative, or zero as we saw in pseudo-Riemannian spaces, like the Minkowsky spacetime. For now we shall consider what normally refer as Riemannian where the metric

$$ds^2 = g_{cd} (x) dx^c dx^d,$$  \hspace{1cm} (17.135)

is positive. It is not possible, in general, to find a coordinate transformation that transforms the metric into Euclidean form

$$ds^2 = (dx^1)^2 + (dx^2)^2 + \ldots + (dx^N)^2 = \delta_{ab} dx^a dx^b,$$  \hspace{1cm} (17.136)

for all points on the manifold. However, it is possible to find coordinates $x^a$ such that at the point $P$ the new metric functions $g'_{ab}(x)$ satisfy the conditions

$$g'_{ab}(x^a_P) = \delta_{ab},$$  \hspace{1cm} (17.137)

$$\frac{\partial g'_{ab}(x^a)}{\partial x^c} \bigg|_{x^a_P} = 0.$$  \hspace{1cm} (17.138)
Thus in the neighborhood of point $P$, we have
\[ g'_{ab}(x') = \delta_{ab} + O\left[(x' - x'_p)^2\right] \]

We recall from Mathematical methods for any function, $g(x)$, that is differentiable for all values of $x$ in the specified domain,
\[ \frac{d^n g(x)}{dx^n} \] exists for all $n \geq 0$ and $x \in \mathbb{R}$,

one can write the series expansion about $x_p$ in the domain (Taylor series)
\[ g(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{d^n g(x')}{dx^n} \bigg|_{x=x_p} (x - x_p)^n. \quad (17.139) \]

Considering only up to the second order terms,
\[ g(x) = g(x) + \frac{dg(x')}{dx'} \bigg|_{x=x_p} (x - x_p) + \frac{1}{2!} \frac{d^2 g(x')}{dx'^2} \bigg|_{x=x_p} (x - x_p)^2 + \ldots \quad (17.140) \]

For a function of two variable $g(x) = g(x^1, x^2)$, this becomes
\[ g(x^1, x^2) = g(x^1_p, x^2_p) + \frac{\partial g(x^1, x^2)}{\partial x^1} \bigg|_{x^1=x^1_p, x^2=x^2_p} (x^1 - x^1_p)
\]
\[ + \frac{\partial g(x^1, x^2)}{\partial x^2} \bigg|_{x^1=x^1_p, x^2=x^2_p} (x^2 - x^2_p) + \frac{1}{2!} \left\{ \frac{\partial^2 g(x^1, x^2)}{\partial x^1 \partial x^1} \bigg|_{x^1=x^1_p, x^2=x^2_p} (x^1 - x^1_p)^2
\]
\[ + 2 \frac{\partial^2 g(x^1, x^2)}{\partial x^1 \partial x^2} \bigg|_{x^1=x^1_p, x^2=x^2_p} (x^1 - x^1_p)(x^2 - x^2_p)
\]
\[ + \frac{\partial^2 g(x^1, x^2)}{\partial x^2 \partial x^2} \bigg|_{x^1=x^1_p, x^2=x^2_p} (x^2 - x^2_p)^2 \right\} + \ldots \quad (17.141) \]

From this expression you can imagine how it gets nasty for a function of $N$ variables, like the metric, $g_{ab}(x) = g_{ab}(x^1, x^2, x^3, \ldots x^N)$.

**Example 2.8** Let’s reconsider the metric for 2D sphere embedded in a 3D manifold in Cartesian coordinates
\[ ds^2 = (dx^1)^2 + (dx^2)^2 + \left( \frac{x^1 dx^1 + x^2 dx^2}{\sqrt{a^2 - (x^1)^2 - (x^2)^2}} \right)^2. \quad (17.142) \]

which can be put in the form
\[ ds^2 = \left[ 1 + \frac{(x^1)^2}{a^2 - (x^1)^2 - (x^2)^2} \right] (dx^1)^2 + \left[ 1 + \frac{(x^2)^2}{a^2 - (x^1)^2 - (x^2)^2} \right] (dx^2)^2
\]
\[ + \frac{2x^1 x^2}{a^2 - (x^1)^2 - (x^2)^2} dx^1 dx^2. \quad (17.143) \]
where \((x^1, x^2, x^3)\) corresponds to the usual Cartesian coordinates \((x, y, z)\). If one pic a point \(P\) with coordinates \((x^1_p, x^2_p)\) so that one may introduce the transformation defined by
\[
x^1 = x^1 - x^1_p, x^2 = x^2 - x^2_p \Rightarrow dx^1 = dx^1, dx^2 = dx^2
\]
and express the metric as
\[
ds^2 = g_{11} \left(dx^1\right)^2 + g_{22} \left(dx^2\right)^2 + g_{12} dx^1 dx^2.
\]

where
\[
g_{11} \left(x^1, x^2\right) = 1 + \frac{(x^1 - x^1_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2},
\]
\[
g_{22} \left(x^1, x^2\right) = 1 + \frac{(x^2 - x^2_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2},
\]
\[
g_{12} \left(x^1, x^2\right) = \frac{2 (x^1 - x^1_p) (x^2 - x^2_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2}
\]

Show that the metric is Locally Cartesian at point \(P\).

**Solution:** For a locally Cartesian, one must be able to show that at point \(P\)
\[
g'_{ab} \left(x^a_p\right) = \delta_{ab}, \frac{\partial g'_{ab} \left(x^a\right)}{\partial x^c} \bigg|_{x^a_p} = 0. \quad (17.144)
\]

We note that for the 2D sphere in the 3D manifold, using the metric we find
\[
g_{11} \left(x^1_p, x^2_p\right) = g_{22} \left(x^1_p, x^2_p\right) = 1, g_{12} \left(x^1_p, x^2_p\right) = 0
\]
and
\[
\frac{\partial g_{11} \left(x^1, x^2\right)}{\partial x^1} \bigg|_{x^1_p, x^2_p} = \frac{\partial}{\partial x^1} \left[ \frac{(x^1 - x^1_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_p
\]
\[
= \left[ \frac{2 (x^1 - x^1_p)}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right] \left[ \frac{2 (x^1 - x^1_p)^3}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_p
\]
\[
= 0
\]

Similarly one can easily show that
\[
\frac{\partial g_{11} \left(x^1, x^2\right)}{\partial x^2} \bigg|_{x^1_p, x^2_p} = \frac{\partial}{\partial x^2} \left[ \frac{(x^1 - x^1_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_p
\]
\[
= \left[ \frac{2 (x^1 - x^1_p)^2 (x^2 - x^2_p)}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_p = 0
\]
17.7. LOCAL CARTESIAN COORDINATES AND TANGENT SPACE

The line element is Euclidean. That means the line element is
\[ ds^2 = (dx^1)^2 + (dx^2)^2 + \ldots (dx^N)^2 \]

this space is called the tangent space, \( T_p \). You can see in Fig. 17.10, three different two-dimensional Tangent spaces (shown in red with green rectangular grids) at three different points on a 2D surface embedded in a 3D manifold. Fig. 17.10 In each of these tangent spaces in Fig. 17.10, note that

\[ ds^2 = (dx^1)^2 + (dx^2)^2. \]  

as you can see from the rectangular shape of the grids. We can show this quantitatively using the function that defines the surface. The surface shown in Fig. 17.10 is defined by the function

\[ x^3 (x^1, x^2) = \sin^2(x^1) + \cos(x^2) \]

which one may write as

\[ g (x^1, x^2, x^3) = \sin^2(x^1) + \cos(x^2) - x^3 = 0 \]

\[
\frac{\partial g_22 (x^1, x^2)}{\partial x^1} \bigg|_{x^1_p, x^2_p} = \frac{\partial}{\partial x^1} \left[ \frac{(x^2 - x^2_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_{x^1_p, x^2_p} \\
= \left[ \frac{2 (x^2 - x^2_p)(x^2 - x^2_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_{x^1_p, x^2_p} = 0 \quad (17.145)
\]

\[
\frac{\partial g_22 (x^1, x^2)}{\partial x^2} \bigg|_{x^1_p, x^2_p} = \frac{\partial}{\partial x^2} \left[ \frac{(x^2 - x^2_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_{x^1_p, x^2_p} \\
= \left[ \frac{2 (x^2 - x^2_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_{x^1_p, x^2_p} = 0 \quad (17.146)
\]

\[
\frac{\partial g_{12} (x^1, x^2)}{\partial x^1} \bigg|_{x^1_p, x^2_p} = \frac{\partial}{\partial x^1} \left[ \frac{2 (x^1 - x^1_p)(x^2 - x^2_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_{x^1_p, x^2_p} = 0 \quad (17.147)
\]

\[
\frac{\partial g_{12} (x^1, x^2)}{\partial x^2} \bigg|_{x^1_p, x^2_p} = \frac{\partial}{\partial x^2} \left[ \frac{2 (x^1 - x^1_p)(x^2 - x^2_p)^2}{a^2 - (x^1 - x^1_p)^2 - (x^2 - x^2_p)^2} \right]_{x^1_p, x^2_p} \quad (17.148)
\]

Tangent Space to Manifolds: for an arbitrary point \( P \) in an \( N \)-dimensional Riemannian manifold we can find a space that consist of coordinates \( (x^a) \) such that in the Neighborhood of \( P \) the line element is Euclidean. That means the line element is

\[ ds^2 = (dx^1)^2 + (dx^2)^2 + \ldots (dx^N)^2 \]
so that

\[ dg = \frac{\partial g}{\partial x^1} dx^1 + \frac{\partial g}{\partial x^2} dx^2 + \frac{\partial g}{\partial x^3} dx^3 = 0 \]

\[
\Rightarrow dg = \left( \frac{\partial g}{\partial x^1} \hat{x} + \frac{\partial g}{\partial x^2} \hat{y} + \frac{\partial g}{\partial x^3} \hat{z} \right) \cdot (dx^1 \hat{x} + dx^2 \hat{y} + dx^3 \hat{z}) = 0
\]

\[
dg = \vec{A} (x^1, x^2, x^3) \cdot d\vec{r} (x^1, x^2, x^3) = 0,
\]

(17.151)

where

\[
\vec{A} (x^1, x^2, x^3) = \frac{\partial g}{\partial x^1} \hat{x} + \frac{\partial g}{\partial x^2} \hat{y} + \frac{\partial g}{\partial x^3} \hat{z}
\]

\[
\frac{\partial g}{\partial x^1} = 2 \sin(x^1) \cos(x^1), \quad \frac{\partial g}{\partial x^2} = -\sin(x^2), \quad \frac{\partial g}{\partial x^3} = -1
\]

(17.152)

Note Eq. (17.151) shows that the vector \( \vec{A} (x^1, x^2, x^3) \) is normal to the surface at the point with coordinates \((x^1, x^2, x^3)\). Now let’s pick a point, \( P \), on the surface with coordinates \( \vec{r}_p = (x^1_p, x^2_p, x^3_p) \) and another neighboring point \( Q \) with coordinates \( \vec{r} = (x^1, x^2, x^3) \) that are at the same plane that is tangent to the surface at point \( P \), then we have for the vector on this plane given by

\[
\Delta \vec{r} = \vec{r} - \vec{r}_p = (x^1 - x^1_p) \hat{x} + (x^2 - x^2_p) \hat{y} + (x^3 - x^3_p) \hat{z}
\]

(17.153)

and the vector normal to this tangent plane at point \( P \) is given by

\[
\vec{A} (x^1_p, x^2_p, x^3_p) = \left[ \frac{\partial g}{\partial x^1} \hat{x} + \frac{\partial g}{\partial x^2} \hat{y} + \frac{\partial g}{\partial x^3} \hat{z} \right]_{x^1_p, x^2_p, x^3_p} = m_1 \hat{x} + m_2 \hat{y} - \hat{z}
\]

(17.154)
where
\[ m_1 = 2 \sin(x_p^1) \cos(x_p^1), \quad m_2 = -\sin(x_p^2) \] (17.155)
The equation of the tangent plane at this point is determined by
\[ g_0^{cd}(x_p) = g_{ab}(x_p) \frac{\partial x^a}{\partial x^c} \frac{\partial x^b}{\partial x^d}. \]
(17.160)
Then at point, \( P \), if the coordinates are \( x'_p = x_p^1, x_p^2, x_p^3, \ldots x_p^N \),
\[ g'_{cd}(x_p) = g_{ab}(x_p) \frac{\partial x^a}{\partial x^c} \frac{\partial x^b}{\partial x^d} \bigg|_{x'_p = x_p}, \]
(17.161)
and if the transformation leads to Local Cartesian, we must have
\[ g'_{cd}(x_p) = \delta_{cd} \lambda_c, \quad \left. \frac{\partial g'_{ab}(x'_p)}{\partial x^c} \right|_{x'_p = x_p} = 0, \]
(17.162)
for pseudo-Riemannian manifolds in general. In fact \( \lambda_c = 1 \) for Local Cartesian in a Riemannian manifolds. Therefore, the transformation metric
\[ g'_{cd}(x') = g_{ab}(x') \frac{\partial x^a}{\partial x^c} \frac{\partial x^b}{\partial x^d}, \]
(17.163)
for local Cartesian at point $P$, one can write

$$
g'_{cd}(x_p) = \delta_{cd} \gamma_{e} = g_{ab}(x') \left(\frac{\partial x^e}{\partial x'^d}\right)_{x_p} \left(\frac{\partial x^b}{\partial x'^a}\right)_{x_p} \quad \begin{align*}
\Rightarrow g'_{cd}(x_p) &= \left(\frac{\partial x^a}{\partial x'^c}\right)_{x_p} g_{ab}(x_p) \left(\frac{\partial x^b}{\partial x'^d}\right)_{x_p}. \quad (17.165)
\end{align*}
$$

This can be put using matrices as

$$
G' = X^T G X. \quad (17.166)
$$

One can easily show this by considering, $N = 2$, and the matrices

$$
G' = \begin{bmatrix}
\lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix},
G = \begin{bmatrix}
g_{11} \left(x^1_p, x^2_p\right) & g_{12} \left(x^1_p, x^2_p\right) \\
g_{21} \left(x^1_p, x^2_p\right) & g_{22} \left(x^1_p, x^2_p\right)
\end{bmatrix}
$$

$$
X = \begin{bmatrix}
\frac{\partial x^1(x^1, x^2)}{\partial x^1} \left|x^1_p, x^2_p\right| & \frac{\partial x^1(x^1, x^2)}{\partial x^2} \left|x^1_p, x^2_p\right| \\
\frac{\partial x^2(x^1, x^2)}{\partial x^1} \left|x^1_p, x^2_p\right| & \frac{\partial x^2(x^1, x^2)}{\partial x^2} \left|x^1_p, x^2_p\right|
\end{bmatrix},
$$

$$
X^T = \begin{bmatrix}
\frac{\partial x^1(x^1, x^2)}{\partial x^1} \left|x^1_p, x^2_p\right| & \frac{\partial x^2(x^1, x^2)}{\partial x^1} \left|x^1_p, x^2_p\right| \\
\frac{\partial x^1(x^1, x^2)}{\partial x^2} \left|x^1_p, x^2_p\right| & \frac{\partial x^2(x^1, x^2)}{\partial x^2} \left|x^1_p, x^2_p\right|
\end{bmatrix} \quad (17.167)
$$

Since the matrix $G'$ is a diagonal matrix, the transformation is a similarity transformation (Mathematical Methods II). This means $X$ forms the eigenvector matrix and $G'$ forms the eigenvalue matrix and

$$
X^T = X^{-1} \Rightarrow G' = X^{-1} G X. \quad (17.168)
$$

Using the inverse transformation matrix, $X'$, and the corresponding inverse matrix, $X'^{-1}$, one must recover the matrix $G$,

$$
G = X'^{-1} G' X' = X'^{-1} X^{-1} G X X' \quad (17.169)
$$

Therefore we must have

$$
XX' = I \Rightarrow X' = X^{-1} \quad (17.170)
$$

This can be true if the transformation from $x^a \rightarrow x'^a$ is linear. The transformation matrix, $X$, with element $X^a_b$, at point $P$ each element must be a constant and we can write

$$
x'^a = X^a_b x^b. \quad (17.171)
$$
Since the metric tensor is symmetric it can be diagonalized by a similarity transformation provided we chose the columns of \( X \) to be the normalized eigenvectors of the matrix \( G \). This means
\[
G' = \begin{bmatrix}
\lambda_1 & 0 & \ldots & 0 \\
0 & \lambda_2 & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & \lambda_N
\end{bmatrix}, \tag{17.172}
\]
where \( \lambda_1, \lambda_2, \ldots, \lambda_N \) are the eigenvalues of the matrix \( G \). The metric
\[
d s^2 = g_{cd} (x') \, dx^c dx^d, \tag{17.173}
\]
is positive for strictly Riemannian. This means \( g_{cd} (x) \) must be positive definite and the eigenvalues, \( \lambda_2 \), for \( G \) are also positive definite. On the other hand for Pseudo Riemannian since the metric can be negative the eigenvalues can be negative. Now if we scale the coordinates \( x'^a \) by these eigenvalues (i.e. \( x'^a \to x'^a / \sqrt{|\lambda_a|} \)), for the metric tensor, \( G' \), we can easily show
\[
G' = \begin{bmatrix}
\pm 1 & 0 & \ldots & 0 \\
0 & \pm 1 & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & \pm 1
\end{bmatrix}. \tag{17.174}
\]
Thus, at any arbitrary point, \( P \), in a pseudo-Riemannian manifold, it is always possible to find a coordinate system \( x'^a \) such that in the neighborhood of \( P \) we have
\[
g'_{ab} (x') = \eta_{ab} + O \left( (x' - x'_p)^2 \right), \tag{17.175}
\]
where \( \eta_{ab} = \text{diag}(\pm 1, \pm 1, \ldots, \pm 1) \). The number of positive entries \( (N_+) \) minus the number of negative entries \( (N_-) \) in \( \eta_{ab} \) is called the \textit{signature} of the manifold. For example for the Minkowski spacetime manifold where the metric is given by
\[
d s^2 = d(ct)^2 - dx^2 - dy^2 - dz^2, \tag{17.176}
\]
the signature is \(-2\).

\textit{N-dimensional volume without a constraint:} In an N-dimensional (pseudo) Riemannian manifold with orthogonal coordinates system where the metric tensor is diagonal, the full N-dimensional volume element \( d^N V \) is
\[
d^N V = \sqrt{|g|} dx^1 dx^2 dx^3 \ldots dx^N \tag{17.178}
\]
where $|g|$ is the determinant of the matrix

$$G = [g_{ab}] = \begin{bmatrix} g_{11} & 0 & \ldots & 0 \\ 0 & g_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & g_{NN} \end{bmatrix}$$

(17.179)
Chapter 18

Vector Calculus on manifolds

18.1 The tangent vector

The tangent vector, \( \vec{t} \), at point \( p \) on a Manifold is the vector that lies in the tangent space, \( T_p \), at that point, \( p_1 \), and is given by

\[
\vec{t} = \lim_{\delta u \to 0} \frac{\delta \vec{s}}{\delta u} = \lim_{\delta u \to 0} \frac{\vec{s}(u + \delta u) - \vec{s}(u)}{\delta u},
\]

where \( \delta s \) is the infinitesimal separation vector between the point \( P \) with coor-

dinate \( (u, s(u)) \) and some nearby point \( Q \) with coordinate \( (u + \delta u, s(u + \delta u)) \)
on the curve, \( C \), in the manifold corresponding (see Fig.??).

Figure 18.1: A tangent vector at point \( P \) in the tangent space.

521
18.2 The basis vectors

At each point \( P \) on a manifold we can define a set of linearly independent basis vectors, \( \mathbf{e}_a(x) \), for the tangent space, \( T_p \). The number of the basis vectors is equal to the dimension of \( T_p \). Any vector field at point \( P \), \( \mathbf{v}(x) \), can then be expressed as linear combination of these basis vectors

\[
\mathbf{v}(x) = v^a(x) \mathbf{e}_a(x),
\]

where \( v^a(x) \) are known as the \text{contravariant components} of the vector field, \( \mathbf{v}(x) \), in the basis \( \mathbf{e}_a \). For any set of basis vectors we can define another set of basis vectors known as the \text{dual basis vectors}, \( \mathbf{\tilde{e}}_a(x) \), defined by

\[
\mathbf{\tilde{e}}_a(x) \cdot \mathbf{e}_b(x) = \delta^a_b. \tag{18.3}
\]

The dual basis vectors, \( \mathbf{\tilde{e}}_a \), and the basis vectors, \( \mathbf{e}_b \), form a reciprocal system of vectors. The local vector field \( \mathbf{v}(x) \) can also be expressed in terms of the \text{dual basis vectors} as

\[
\mathbf{v}(x) = v_a(x) \mathbf{\tilde{e}}_a(x), \tag{18.4}
\]

where \( v_a(x) \) are known as the \text{covariant components} of the vector field, \( \mathbf{v}(x) \), in the dual basis vectors, \( \mathbf{\tilde{e}}_a(x) \). The controvariant and covariant components of the vector field can be determined using Eq. (18.3)

\[
\mathbf{v}(x) \cdot \mathbf{\tilde{e}}^b(x) = v^a(x) \mathbf{\tilde{e}}_a(x) \cdot \mathbf{\tilde{e}}^b(x) = v^a(x) \delta^a_b = v^b(x). \tag{18.5}
\]

Similarly

\[
v_a(x) = \mathbf{v}(x) \cdot \mathbf{e}_a(x). \tag{18.6}
\]

The \text{coordinate basis vectors}: in any particular coordinate system \( x^a \), we can define at every point \( P \) of the manifold a set of \( N \) coordinate basis vectors

\[
\mathbf{e}_a = \lim_{\delta x^a \to 0} \frac{\delta \mathbf{s}}{\delta x^a}, \tag{18.7}
\]

where \( \delta \mathbf{s} \) is the infinitesimal separation vector between point \( P \) and some nearby point \( Q \) with coordinate separation \( \delta x^a \) from \( P \). For example, \( \mathbf{e}_a \) is the tangent vector to the \( x^a \) coordinate curve at the point \( P \). As an example we reconsider the 2D (blue) surface embedded in a 3D Manifold shown in Fig. 18.2. We recall that this surface is defined by

\[
x^3(x^1, x^2) = \sin^2(x^1) + \cos(x^2) \tag{18.8}
\]

where \( x^1 \), \( x^2 \), and \( x^3 \) are the usual Cartesian coordinates \( (x,y) \). The tangent plane at point \( P \) is given by

\[
x^3(x^1, x^2) = 2 \sin(x^1_p) \cos(x^1_p) (x^1 - x^1_p) - \sin(x^2_p) (x^2 - x^2_p) + x^3_p. \tag{18.9}
\]

where \( (x^1_p, x^2_p, x^3_p) \) are the coordinates for point \( P \). We note that a point on the surface can defined by a vector

\[
\mathbf{s} = x^1 \mathbf{x} + x^2 \mathbf{y} + (\sin^2(x^1_p) + \cos(x^2)) \mathbf{z} \tag{18.10}
\]
so that
\[ \delta \hat{s} = \delta x^1 \hat{x} + \delta x^2 \hat{y} + \left( 2 \sin(x^1) \cos(x^1) \delta x^1 - \sin(x^2) \delta x^2 \right) \hat{z} \]  
(18.11)

The two basis vectors, \( \hat{e}_1 \) and \( \hat{e}_2 \), at point \( P \), in Fig. 18.2 are given by
\[ \hat{e}_1 = \lim_{\delta x^1 \to 0} \frac{\delta \hat{s}}{\delta x^1} = \lim_{\delta x^1 \to 0} \left[ \hat{x} + \delta x^2 \delta x^1 \hat{y} + \left( 2 \sin(x^1) \cos(x^1) - \sin(x^2) \frac{\delta x^2}{\delta x^1} \right) \hat{z} \right] \]  
(18.12)
\[ \hat{e}_2 = \lim_{\delta x^2 \to 0} \frac{\delta \hat{s}}{\delta x^2} = \lim_{\delta x^2 \to 0} \left[ \frac{\delta x^1}{\delta x^2} \hat{x} + \hat{y} + \left( 2 \sin(x^1) \frac{\delta x^1}{\delta x^2} - \sin(x^2) \right) \hat{z} \right] \]  
(18.13)

and noting that
\[ \frac{\delta x^1}{\delta x^2} = \frac{\delta x^2}{\delta x^1} = 0 \]
we find
\[ \hat{e}_1 = \hat{x} + 2 \sin(x^1) \cos(x^1) \hat{z}, \hat{e}_2 = \hat{y} - \sin(x^2) \hat{z} \]  
(18.14)

### 18.3 The metric function and coordinate transformations

**Infinitesimal vector separation**: Consider two points \( P \) and \( Q \) on a manifold with coordinates \( x^a \) and \( x^a + dx^a \), where \( dx^a \) is none zero for all \( a \), then the infinitesimal vector separation between these two points is given by
\[ d\hat{s} = \hat{e}_a (x) dx^a \]  
(18.15)
The metric function—the covariant components: The equation that determines the elements of the metric tensor in the metric

\[ ds^2 = g_{ab}(x) \, dx^a \, dx^b \]  

(18.16)
can be obtained from the inner product of the infinitesimal vector separation. We note that

\[
    ds^2 = d\tilde{s} \cdot d\tilde{s} = \hat{e}_a(x) \, dx^a \cdot \hat{e}_b(x) \, dx^b = \hat{e}_a(x) \cdot \hat{e}_b(x) \, dx^a \, dx^b \\
    \Rightarrow \quad ds^2 = g_{ab}(x) \, dx^a \, dx^b, \tag{18.17}
\]

where

\[ g_{ab}(x) = \hat{e}_a(x) \cdot \hat{e}_b(x) \]
is the metric function

**Example 3.1**  Let’s reconsider the 2D sphere of radius, \( a \), in a 3D Euclidean space. For an origin set at the north pole of the sphere, a point \( P \) is described by the vector

\[ \tilde{s} = \rho \cos(\varphi) \hat{x} + \rho \sin(\varphi) \hat{y} + \sqrt{a^2 - \rho^2} \hat{z} \]

Figure 18.3: 2D sphere of radius \( a \) embedded in 3D Euclidean space. The origin of the Cartesian coordinates \((x^1, x^2, x^3)\) [or \((x, y, z)\)] is set at the north pole. \((x^0 = \rho, x'^0 = \varphi)\).

(a) Find the basis vectors \( \hat{e}'_1 \) and \( \hat{e}'_2 \) in the tangent space at point \( P \).
18.3. THE METRIC FUNCTION AND COORDINATE TRANSFORMATIONS

(b) Re-develop the metric elements for a 2D sphere from the basis vectors.

Solution:

(a) We note that the tangent vector connecting point $P$ with coordinate $(\rho, \varphi)$ with its neighboring point $Q$ with coordinates $(\rho + \delta \rho, \varphi + \delta \varphi)$ (See Fig. 18.4) is expressible as

$$
\delta \vec{s} = \delta (\rho \cos(\varphi)) \hat{x} + \delta (\rho \sin(\varphi)) \hat{y} + \delta \left(\frac{\sqrt{a^2 - \rho^2}}{a^2} \hat{z}\right)
$$

$$
= [\cos(\varphi) \delta \rho - \rho \sin(\varphi) \delta \varphi] \hat{x} + [\sin(\varphi) \delta \rho + \rho \cos(\varphi) \delta \varphi] \hat{y}
$$

$$
- \frac{\rho \delta \rho}{\sqrt{a^2 - \rho^2}} \hat{z}
$$

(18.18)

$$
= \left[\cos(\varphi) \hat{x} + \sin(\varphi) \hat{y} - \frac{\rho}{\sqrt{a^2 - \rho^2}} \hat{z}\right] \delta \rho + [-\rho \sin(\varphi) \hat{x} + \rho \cos(\varphi) \hat{y}] \delta \varphi
$$

(18.19)

There follows that

$$
\hat{e}_\rho = \lim_{\delta \rho \to 0} \frac{\delta \vec{s}}{\delta \rho}
$$

$$
= \left[\cos(\varphi) \hat{x} + \sin(\varphi) \hat{y} - \frac{\rho}{\sqrt{a^2 - \rho^2}} \hat{z}\right] \delta \rho + [-\rho \sin(\varphi) \hat{x} + \rho \cos(\varphi) \hat{y}] \frac{\delta \varphi}{\delta \rho}.
$$

(18.20)

and

$$
\hat{e}_\varphi = \lim_{\delta \varphi \to 0} \frac{\delta \vec{s}}{\delta \varphi}
$$

$$
= \left[\cos(\varphi) \hat{x} + \sin(\varphi) \hat{y} - \frac{\rho}{\sqrt{a^2 - \rho^2}} \hat{z}\right] \frac{\delta \rho}{\delta \varphi} + [-\rho \sin(\varphi) \hat{x} + \rho \cos(\varphi) \hat{y}] \frac{\delta \varphi}{\delta \varphi}.
$$

(18.21)

Noting that

$$
\frac{\delta \rho}{\delta \varphi} = \frac{\delta \rho}{\delta \varphi} = 0
$$

(18.22)

we find

$$
\hat{e}_\rho = \cos(\varphi) \hat{x} + \sin(\varphi) \hat{y} - \frac{\rho}{\sqrt{a^2 - \rho^2}} \hat{z}.
$$

(18.23)

and

$$
\hat{e}_\varphi = -\rho \sin(\varphi) \hat{x} + \rho \cos(\varphi) \hat{y}.
$$

(18.24)
Figure 18.4: The basis vectors \((\hat{e}_\rho, \hat{e}_\varphi)\) at a point \(P\) in the tangent space of a 2D sphere embedded in a 3D Euclidean space.

(b) The metric elements can be determined using the basis vectors

\[
g_{ab}(x) = \hat{e}_a(x) \cdot \hat{e}_b(x), \tag{18.25}
\]

and one finds

\[
g_{11} = g_{\rho\rho} = \hat{e}_\rho \cdot \hat{e}_\rho = 1 - \frac{\rho^2}{a^2 - \rho^2} = \frac{a^2}{a^2 - \rho^2},
\]

\[
g_{22} = g_{\varphi\varphi} = \hat{e}_\varphi \cdot \hat{e}_\varphi = \rho^2, g_{12} = g_{21} = 0 \tag{18.26}
\]

**Homework:** Consider the 3D sphere of radius, \(a\), in a 4D Euclidean space. A point \(P\) on this 3D sphere is described by the vector

\[
\vec{s} = r \sin(\theta) \cos(\varphi) \hat{x} + r \sin(\theta) \sin(\varphi) \hat{y} + r \cos(\theta) \hat{z} + \sqrt{a^2 - r^2} \hat{w} \tag{18.27}
\]

(a) Find the basis vectors \(\hat{e}_r, \hat{e}_\theta,\) and \(\hat{e}_\varphi\) in the tangent space at point \(P\).

(b) Re-derive the metric elements for a 3D sphere from the basis vectors.

The **metric function**—the covariant components: noting that the infinitesimal distance between point \(P\) and \(Q\) can be expressed using the infinitesimal vector separation

\[
d\vec{s} = \hat{e}_a(x) dx^a \tag{18.28}
\]
we have

$$\hat{e}^b(x) \cdot d\vec{s} = \hat{e}^b(x) \cdot \hat{e}_a(x) \, dx^a = \delta_a^b \, dx^a = dx^b.$$  \hspace{1cm} (18.29)

Expressing the infinitesimal vector separation using the dual coordinates basis vectors as

$$d\vec{s} = \hat{e}^a(x) \, dx^a$$  \hspace{1cm} (18.30)

we also find

$$\hat{e}_b(x) \cdot d\vec{s} = \hat{e}_b(x) \cdot \hat{e}^a(x) \, dx^a = \delta^b_a \, dx^a = dx^b.$$  \hspace{1cm} (18.31)

Thus the metric

$$d\vec{s}^2 = d\vec{s} \cdot d\vec{s} = \hat{e}^a(x) \, dx^a \cdot \hat{e}^b(x) \, dx^b = g^{ab}(x) \, dx^a \, dx^b,$$  \hspace{1cm} (18.32)

where the controvariant components of the metric tensor can be defined as

$$g^{ab}(x) = \hat{e}^a(x) \cdot \hat{e}^b(x)$$  \hspace{1cm} (18.33)

**Orthonormal basis vector:** at a point on a manifold an orthonormal basis vectors are defined by

$$\hat{e}_a(x) \cdot \hat{e}_b(x) = \eta_{ab},$$  \hspace{1cm} (18.34)

where

$$[\eta_{ab}] = \begin{bmatrix} \pm 1 & 0 & \cdots & 0 \\ 0 & \pm 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm 1 \end{bmatrix}.$$  \hspace{1cm} (18.35)

or in short $[\eta_{ab}] = diag(\pm 1, \pm 1, \ldots, \pm 1)$ is the Cartesian line element of the tangent space, $T_p$, and depends on the signature of the pseudo-Riemannian manifold.

**Basis vectors and coordinate transformations:** Suppose we make a coordinate transformation from $x^a$ where the basis vectors are $\hat{e}_a$ to a new another coordinates $\tilde{x}^a$ we want to find how the new basis vectors are transformed into, $\hat{e}_c'$. If we consider point $P$ and another point $Q$ at an infinitesimal distance away from point $P$. The infinitesimal displacement, $d\vec{s}'$, between these points is independent of the coordinate transformation and we must have

$$d\vec{s}' = \hat{e}_a \, dx^a = \hat{e}_c' \, d\tilde{x}^c.$$  \hspace{1cm} (18.36)

Using

$$dx^a = \frac{\partial x^a}{\partial \tilde{x}^b} \, d\tilde{x}^b,$$  \hspace{1cm} (18.37)

we have

$$d\vec{s}' = \hat{e}_a \, dx^a = \frac{\partial x^a}{\partial \tilde{x}^b} \hat{e}_a \, d\tilde{x}^b = \hat{e}_c' \, d\tilde{x}^c.$$  \hspace{1cm} (18.38)

Noting that

$$\hat{e}_b \cdot \hat{e}^b = 1$$
we have
\[
\frac{\partial x^a}{\partial x'^b} \hat{e}_a dx^b = (\hat{e}'_c \cdot \hat{e}^b) \hat{e}'_c dx^c = (\hat{e}'_c \cdot \hat{e}^b) \hat{e}'_d dx^c = \hat{e}'_d \delta^b_c dx^c \tag{18.39}
\]
\[
\frac{\partial x^a}{\partial x'^b} \hat{e}_a dx^b = \hat{e}'_b dx^b. \tag{18.40}
\]
so that
\[
\hat{e}'_b = \frac{\partial x^a}{\partial x'^b} \hat{e}_a. \tag{18.41}
\]

**Homework:** Show that for the dual basis vector
\[
\hat{e}^0_a = \frac{\partial x^a}{\partial x^0} \hat{e}_c. \tag{18.42}
\]

**Components of a vector in coordinate transformations:** In coordinate transformations the vector components are different but the vector itself is unchanged. Suppose the vector, $\vec{v}$, is a vector at point $P$ in the $x^a$ coordinate system and $\vec{v}'$ is the vector in the $x'^a$ coordinates at the same point on the manifold. These vectors may be expressed in terms of the basis vectors in the two coordinates differently
\[
\vec{v} = v^a \hat{e}_a; \quad \vec{v}' = v'^a \hat{e}'_a \tag{18.43}
\]
or in terms of the dual basis vectors
\[
\vec{v} = v_a \hat{e}^a; \quad \vec{v}' = v'^a \hat{e}'^a \tag{18.44}
\]
But the vector is the same since it describes a geometrical entity that is independent of the coordinate system. Therefore, we must have
\[
\vec{v} = v^a \hat{e}_a; \quad \vec{v}' = v'^a \hat{e}'_a \tag{18.45}
\]
so that taking the inner product of $\vec{v}$ and $\hat{e}'_b$, we can write
\[
v^a \hat{e}'_b \cdot \hat{e}_a = v'^a \hat{e}'_b \cdot \hat{e}'_a \Rightarrow v'^a \hat{e}'_b = v^a \hat{e}^b \cdot \hat{e}_a \Rightarrow v'^b = v^a \hat{e}^b \cdot \hat{e}_a. \tag{18.46}
\]
Applying the relation in Eq. (18.42), one can write
\[
\hat{e}'_b = \frac{\partial x'^b}{\partial x^c} \hat{e}_c, \tag{18.47}
\]
so that
\[
v'^b = v^a \frac{\partial x'^b}{\partial x^c} \hat{e}_c \cdot \hat{e}_a = v^a \frac{\partial x'^b}{\partial x^c} \delta^c_a = \frac{\partial x'^b}{\partial x^c} v^a. \tag{18.48}
\]
**Homework:** Show that for the covariant components of a vector transformed by the equation
\[
v'_b = \frac{\partial x^a}{\partial x'^b} v_a. \tag{18.49}
\]
18.3. THE METRIC FUNCTION AND COORDINATE TRANSFORMATIONS

18.3.1 Raising and lowering vector indices

The scalar product: The scalar product of two vectors at a point, P, on a manifold

\[ \vec{v} = v^a \hat{e}_a. \]  

(18.50)

and

\[ \vec{w} = w^b \hat{e}_b. \]  

(18.51)

is given by

\[ \vec{v} \cdot \vec{w} = v^a \hat{e}_a \cdot w^b \hat{e}_b = g_{ab} v^a w^b. \]  

(18.52)

where

\[ g_{ab} = \hat{e}_a \cdot \hat{e}_b. \]  

(18.53)

is the covariant components of the metric tensor. We can also use the dual basis vectors to express the vectors \( \vec{v}(x) \) and \( \vec{w}(x) \)

\[ \vec{v} = v^a \hat{e}_a. \]  

(18.54)

and

\[ \vec{w} = w^b \hat{e}_b. \]  

(18.55)

so that the inner product becomes

\[ \vec{v} \cdot \vec{w} = v^a \hat{e}_a \cdot w^b \hat{e}_b = g^{ab} v_a w_b. \]  

(18.56)

where

\[ g^{ab} = \hat{e}_a \cdot \hat{e}_b. \]  

(18.57)

We can use the covariant and contravariant components of the vectors to determine the inner products

\[ \vec{v} \cdot \vec{w} = v^a \hat{e}_a \cdot w^b \hat{e}_b = \delta^a_b v^a w_b = v^a w_a \]  

(18.58)

or

\[ \vec{v} \cdot \vec{w} = v_a \hat{e}^a \cdot w^b \hat{e}_b = \hat{e}_a \cdot \hat{e}_b v_a w^b = \delta^a_b v_a w^b = v_a w^a. \]  

(18.59)

Whichever way we determine the inner products we must get the same values. Thus from Eqs. (18.52) and (18.58), we find

\[ \vec{v} \cdot \vec{w} = g_{ab} v^a w^b = v^a w_a \Rightarrow g_{ab} w^b = w_a. \]  

(18.60)

Similarly, from Eqs. (18.56) and (18.59), we find

\[ \vec{v} \cdot \vec{w} = g^{ab} v_a w_b = v_a w^a \Rightarrow g^{ab} w_b = w^a. \]  

(18.61)

From Eq. (18.60), we note that the covariant form of the metric tensor can be used to lower an index and from Eq. (18.61) we also see that the contravariant form of the metric tensor can be used to raise an index. Applying Eqs. (18.60) and (18.61), we can express the basis vectors

\[ \hat{e}_a = g_{ad} \hat{e}^d \]  

(18.62)
and
\[ \hat{e}^c = g^{cb} \hat{e}_b. \]  
(18.63)

Then noting that the inner product
\[ \hat{e}_a \cdot \hat{e}^c = \delta_c^a \Rightarrow g_{ad} \hat{e}^d \cdot g^{cb} \hat{e}_b = g_{ad} g^{cb} \delta_e^d \cdot \hat{e}_b = \delta_c^a \]
\[ \Rightarrow g_{ad} g^{cb} \delta_e^d \cdot \hat{e}_b = g_{ad} g^{cb} \delta_e^d \delta_c^a \]
\[ \Rightarrow g_{ab} g^{cb} = g^{cb} g_{ab} = \delta_c^a. \]  
(18.64)

This means the metric tensor \([g^{ab}]\) with the contravariant components, \(g^{ab}\), is the inverse matrix of the metric tensor \([g_{ab}]\) with the covariant elements, \(g_{ab}\). Thus
\[ G \tilde{G} = \tilde{G} G = I, \]
where \(G = [g_{ab}]\) is the metric tensor and \(\tilde{G} = [g^{ab}]\) is its inverse.

### 18.4 The inner product and null vectors

The scalar product of two vectors at a point on a manifold which can be expressed in four different ways
\[ g_{ab} v^a w^b = v_a w^a = g^{ab} v_a w_b = v^a w_a \]  
(18.65)

Suppose we take the scalar product of vector, \(\vec{v}\), with itself, we have
\[ g_{ab} \vec{v}^a \vec{v}^b = g^{ab} \vec{v}_a \vec{v}_b = \vec{v}^a \vec{v}_a = \vec{v}^a \vec{v}_a \]  
(18.66)

and it can be zero without the vector being actually be a zero vector. We can see this if we recall the pseudo-Riemannian manifold metric
\[ ds^2 = g_{ab}(x) dx^a dx^b \]
which could be zero or negative. To accommodate such kind of vectors we define the length of a vector, \(\vec{v}\), as
\[ v = \sqrt{|g_{ab} v^a v^b|} = \sqrt{|g^{ab} v_a v_b|} = \sqrt{|v^a v_a|} = \sqrt{|v_a v^a|}. \]  
(18.67)

As is the case in pseudo-Riemannian manifold, the length of a vector can be zero without the vector being actually be a zero vector (i.e. \(v_a \neq 0\)). Vectors with length (magnitude) zero with none zero components is known as Null vectors.

**The cosine angle between vectors:** The angle between two non-null vectors at a point on a manifold is defined by:
\[ \cos(\theta) = \frac{v^a w_a}{\sqrt{|v_b v^b|} \sqrt{|w_c w_c|}} \]  
(18.68)

In the pseudo-Riemannian manifold, Eq. (18.68) can lead to \(|\cos(\theta)| > 1\).
18.5 The affine connections

Orthogonal vectors: two vectors

\[ \vec{v} = v^a \hat{e}_a. \] (18.69)

and

\[ \vec{w} = w^b \hat{e}_b. \] (18.70)

are said to be orthogonal when

\[ g^{ab} v_a w^b = g^{ab} v_a w_b = v^a w_a = 0. \] (18.71)

18.5 The affine connections

It is important to know how vectors changes as the coordinate or the parameter that defines the coordinate changes. For example, in Mankowski space time we may be interested in the 4D momentum, \( \vec{P} \), how it changes with time (the proper time, \( \tau \)) so that one can explain the condition for conservation of momentum in general relativity. In such cases for the 4D momentum expressed in terms of its contovariant components as

\[ \vec{P} = p^a \hat{e}_a \] (18.72)

one must be able to determine

\[ \frac{d\vec{P}}{d\tau} = \frac{d}{d\tau} (p^a \hat{e}_a) = \hat{e}_a \frac{dp^a}{d\tau} + p^a \frac{d\hat{e}_a}{d\tau} \] (18.73)

For the coordinates \( x^a = a^a (\tau) \), we have

\[ \frac{d\hat{e}_a}{d\tau} = \hat{e}_a \frac{dx^b}{d\tau} \] (18.74)

so that

\[ \frac{d\vec{P}}{d\tau} = \hat{e}_a \frac{dp^a}{d\tau} + p^a \frac{dx^b}{d\tau} \frac{d\hat{e}_a}{d\tau} \] (18.75)

In order to better understand the origin of the affine connections we shall reconsider the 2D sphere embedded in a 3D manifold. We saw that the tangent space is a plane with basis vectors defined by

\[ \hat{e}_\rho = \cos (\varphi) \hat{x} + \sin (\varphi) \hat{y} - \frac{\rho}{\sqrt{a^2 - \rho^2}} \hat{z}. \] (18.76)

and

\[ \hat{e}_\varphi = -\rho \sin (\varphi) \hat{x} + \rho \cos (\varphi) \hat{y}. \] (18.77)

There follows that

\[ \frac{\partial \hat{e}_\rho}{\partial \varphi} = \frac{\partial \hat{e}_\varphi}{\partial \rho} = -\sin (\varphi) \hat{x} + \cos (\varphi) \hat{y} = \frac{1}{\rho} \hat{e}_\varphi = f_{\rho \varphi} (\rho, \varphi) \hat{e}_\varphi. \] (18.78)
CHAPTER 18. VECTOR CALCULUS ON MANIFOLDS

where

\[ f_{\rho \varphi} (\rho, \varphi) = \frac{1}{\rho}, \]

is a function that connects the change in the basis vectors with respect to the coordinates to the basis vectors. Here we note that we are still in the tangent space. However, if we switch the variables for the derivatives, we find

\[
\frac{\partial \hat{e}_\rho}{\partial \varphi} = -\rho \cos (\varphi) \hat{x} - \rho \sin (\varphi) \hat{y} = -\rho (\cos (\varphi) \hat{x} + \sin (\varphi) \hat{y}) .
\]

that we can not tell whether it belongs to the tangent space or not. In order to find out that we introduce a basis vector normal to the tangent space in terms of the basis vector in the tangent space as

\[
\hat{e}_\perp = \hat{e}_\rho \times \hat{e}_\varphi.
\]

This normal basis vector is found to be

\[
\hat{e}_\perp = \frac{\rho^2}{\sqrt{a^2 - \rho^2}} (\cos (\varphi) \hat{x} + \sin (\varphi) \hat{y}) + \rho \hat{z} .
\]

Combining this relation with the basis vector, \( \hat{e}_\rho \),

\[
\hat{e}_\rho = \cos (\varphi) \hat{x} + \sin (\varphi) \hat{y} - \frac{\rho}{\sqrt{a^2 - \rho^2}} \hat{z} .
\]

one find

\[
\hat{z} = \frac{a^2 - \rho^2}{a^2} \left( \frac{1}{\rho} \hat{e}_\perp - \frac{\rho}{\sqrt{a^2 - \rho^2}} \hat{e}_\rho \right)
\]

and

\[
\cos (\varphi) \hat{x} + \sin (\varphi) \hat{y} = \frac{1}{a} \sqrt{1 - \frac{\rho^2}{a^2}} \hat{e}_\perp - \frac{\rho^2}{a^2} \hat{e}_\rho .
\]

Using these relations, we can then write

\[
\frac{\partial \hat{e}_\rho}{\partial \rho} = -\frac{\rho}{\sqrt{a^2 - \rho^2}} \hat{e}_\rho + \frac{1}{\rho} \hat{e}_\perp, \quad \frac{\partial \hat{e}_\varphi}{\partial \rho} = -\frac{\rho}{a} \sqrt{1 - \frac{\rho^2}{a^2}} \hat{e}_\perp + \frac{\rho^3}{a^2} \hat{e}_\rho .
\]

or

\[
\frac{\partial \hat{e}_\rho}{\partial \rho} = [f_{\rho \rho} (\rho, \varphi)]_\parallel \hat{e}_\rho + [f_{\rho \varphi} (\rho, \varphi)]_\perp \hat{e}_\perp, \quad \frac{\partial \hat{e}_\varphi}{\partial \rho} = [f_{\varphi \rho} (\rho, \varphi)]_\parallel \hat{e}_\rho + [f_{\varphi \varphi} (\rho, \varphi)]_\perp \hat{e}_\perp .
\]

where

\[
[f_{\rho \rho} (\rho, \varphi)]_\parallel = -\frac{\rho}{\sqrt{a^2 - \rho^2}} [f_{\rho \rho} (\rho, \varphi)]_\perp = \frac{1}{\rho}
\]

\[
[f_{\varphi \varphi} (\rho, \varphi)]_\parallel = -\frac{\rho}{a} \sqrt{1 - \frac{\rho^2}{a^2}} [f_{\varphi \varphi} (\rho, \varphi)]_\perp = \frac{\rho^3}{a^2}
\]
These results show that the derivatives for the basis vectors can lead to components that do not belong to the tangent space. Let’s see where these components are located for a specific point for a 2D sphere with radius, \( a = 2 \) units. The point we shall consider has coordinates \((x = 1, y = 0, z = \sqrt{3})\) which is on the surface of the 2D sphere. At this point the two basis vectors \((\hat{e}_\rho, \hat{e}_\varphi)\) and the vector normal to the tangent space, are shown in Fig. 18.5.

\[
\frac{\partial \hat{e}_\rho}{\partial \rho} = -\frac{1}{\sqrt{3}} \hat{e}_\rho + \hat{e}_\perp, \quad \frac{\partial \hat{e}_\varphi}{\partial \varphi} = -\frac{\sqrt{3}}{4} \hat{e}_\perp + \frac{1}{4} \hat{e}_\rho. \tag{18.86}
\]

Therefore, generally for none spherical geometry or none Euclidean manifold, the change in the basis vectors with respect to the coordinate at a point on a manifold (derivative of the basis vector at that point) can have both normal and tangential components to the tangent space. However, since we are confined to the tangent space, for example in the 2D sphere of radius \( a \), we are confined to the tangent space that is a plane and the normal component does not belong to the tangent space at that particular point. Therefore, we shall consider only the projection parallel to the tangent space at point \( P \),

\[
\frac{\partial \hat{e}_a}{\partial x^c} = \left( \lim_{\delta x^c \to 0} \frac{\delta \hat{e}_a}{\delta x^c} \right)_{\| T_P }. \tag{18.87}
\]

Suppose we represent the coefficients resulting from the derivative of the basis vectors that are components that belong to the tangent space at the point (e.g., \( f_{ac}(\rho, \varphi)|_{T_P} \)) by \( \Gamma^b_{ac} \), then we can write

\[
\frac{\partial \hat{e}_a}{\partial x^c} = \Gamma^1_{ac} \hat{e}_1 + \Gamma^2_{ac} \hat{e}_2 + \Gamma^3_{ac} \hat{e}_3 \ldots + \Gamma^N_{ac} \hat{e}_N = \Gamma^b_{ac} \hat{e}_b, \tag{18.88}
\]
where $N$ is the dimension of the tangent space. The $N^3$ coefficients $\Gamma^b_{ac}$ are known collectively as the affine connection or in older textbooks, the Christoffel symbol (of the second kind) at point $P$.

**Homework:**

(1) Find all the elements for the affine connection, $\Gamma^b_{ac}$, for a point on a 2D sphere embedded in a 3D Euclidean space. Note that in the expressions

$$\frac{\partial \epsilon^a}{\partial \theta} = \Gamma^a_{\theta \theta} \epsilon^\theta + \Gamma^a_{\theta \phi} \epsilon^\phi,$$

$$\frac{\partial \epsilon^b}{\partial \theta} = \Gamma^b_{\phi \theta} \epsilon^\theta + \Gamma^b_{\phi \phi} \epsilon^\phi,$$  

$$\frac{\partial \epsilon^c}{\partial \theta} = \Gamma^c_{\phi \theta} \epsilon^\theta + \Gamma^c_{\phi \phi} \epsilon^\phi,$$  

(18.89)

you are going to determine

$$\Gamma^a_{\theta \theta}, \Gamma^a_{\theta \phi}, \Gamma^a_{\phi \theta}, \Gamma^a_{\phi \phi}, \Gamma^b_{\phi \phi}, \Gamma^b_{\phi \phi}, \Gamma^c_{\phi \phi}, \Gamma^c_{\phi \phi}, \Gamma^c_{\phi \phi}, \Gamma^c_{\phi \phi}, \Gamma^c_{\phi \phi}.$$  

(18.90)

Also in this case the origin is at the center of the sphere.

(2) Find all the elements for the affine connection, $\Gamma^b_{ac}$, for a point on a 3D sphere embedded in a 4D Euclidean space. Note that in the expressions

Taking the scalar product of Eq. (18.88) and the dual basis vectors,

$$\epsilon^d \cdot \frac{\partial \epsilon^a}{\partial x^c} = \Gamma^b_{ac} \epsilon^d \cdot \epsilon^b,$$  

(18.91)

applying the properties of the dual basis vectors

$$\epsilon^d \cdot \epsilon^b = \delta^d_b,$$  

(18.92)

we find

$$\epsilon^d \cdot \frac{\partial \epsilon^a}{\partial x^c} = \Gamma^b_{ac} \delta^d_b \Rightarrow \Gamma^d_{ac} = \epsilon^d \cdot \frac{\partial \epsilon^a}{\partial x^c},$$  

(18.93)

Since $d$ is a dummy index, we can write

$$\Gamma^b_{ac} = \epsilon^b \cdot \frac{\partial \epsilon^a}{\partial x^c} = \epsilon^b \cdot \partial_c \epsilon^a.$$  

(18.94)

From now on we will use the notation $\partial_c \epsilon^a$,

$$\partial_c \epsilon^a = \frac{\partial \epsilon^a}{\partial x^c}.$$  

(18.95)

Differentiating Eq. (18.92) with respect to the coordinate, $x^c$, and applying the notation, we find

$$\epsilon^a \cdot \partial_c \epsilon^b + \epsilon^b \cdot \partial_c \epsilon^a = \partial_c \delta^b_a = 0 \Rightarrow \epsilon^b \cdot \partial_c \epsilon^a = -\epsilon^a \cdot \partial_c \epsilon^b,$$  

(18.96)

Using the definition of the derivative of the basis vectors and our notation

$$\partial_c \epsilon^a = \frac{\partial \epsilon^a}{\partial x^c} = \Gamma^b_{ac} \epsilon^b \Rightarrow \epsilon^d \cdot \partial_c \epsilon^a = \Gamma^b_{ac} \epsilon^d \cdot \epsilon^b = \Gamma^d_{ac} \delta^d_b = \Gamma^d_{ac},$$  

(18.97)

$$\Rightarrow \Gamma^d_{ac} = \epsilon^d \cdot \partial_c \epsilon^a.$$  

(18.98)
applying this relation,
\[ \dot{e}_b \cdot \partial c \dot{e}^a = -\dot{e}^a \cdot \partial c \dot{e}_b \Rightarrow \dot{e}_b \cdot \partial c \dot{e}^a = -\Gamma^a_{bc}, \tag{18.99} \]
and noting that
\[ \dot{e}_b \cdot \dot{e}^b = 1 \tag{18.100} \]
we find
\[ \dot{e}_b \cdot \partial c \dot{e}^a = -\Gamma^a_{bc} \dot{e}_b \cdot (\dot{e}^b - \dot{e}^b) \Rightarrow \partial c \dot{e}^a = -\Gamma^a_{bc} \dot{e}^b \tag{18.101} \]

The affine connection under coordinate transformation: Suppose we make the coordinate transformation \( x^a \rightarrow x'^a \), for the affine connection
\[ \Gamma^b_{ac} = \dot{e}^b \cdot \frac{\partial \dot{e}_a}{\partial x'^c}. \tag{18.102} \]
we have
\[ \Gamma'^b_{ac} = \dot{e}'^b \cdot \frac{\partial \dot{e}'_a}{\partial x'^c}. \tag{18.103} \]
so that applying the relations
\[ \dot{e}'_a = \frac{\partial x^f}{\partial x'^a} \dot{e}_f, \dot{e}'^b = \frac{\partial x'^b}{\partial x'^d} \dot{e}_d. \tag{18.104} \]
we may write
\[ \Gamma'^b_{ac} = \frac{\partial x'^b}{\partial x'^d} \dot{e}'_d \cdot \frac{\partial}{\partial x'^c} \left( \frac{\partial x^f}{\partial x'^a} \dot{e}_f \right) = \frac{\partial x'^b}{\partial x'^d} \dot{e}'_d \cdot \left[ \frac{\partial \dot{e}_f}{\partial x'^c} \frac{\partial x^f}{\partial x'^a} + \frac{\partial \dot{e}_f}{\partial x'^c} \frac{\partial^2 x^f}{\partial x'^c \partial x'^a} \right] = \frac{\partial x'^b}{\partial x'^d} \frac{\partial x^f}{\partial x'^a} \dot{e}_d \cdot \frac{\partial \dot{e}_f}{\partial x'^c} + \frac{\partial^2 x^f}{\partial x'^c \partial x'^a} \frac{\partial x'^b}{\partial x'^d} \dot{e}_d \cdot \dot{e}_f. \tag{18.105} \]
Using the chain rule one can write
\[ \frac{\partial \dot{e}_f}{\partial x'^c} = \frac{\partial \dot{e}_f}{\partial x'^g} \frac{\partial x^g}{\partial x'^c} \]
so that Eq. (18.105) becomes
\[ \Gamma'^b_{ac} = \frac{\partial x'^b}{\partial x'^d} \frac{\partial x^f}{\partial x'^a} \frac{\partial x^g}{\partial x'^c} \dot{e}_d \cdot \frac{\partial \dot{e}_f}{\partial x'^g} + \frac{\partial^2 x^f}{\partial x'^c \partial x'^a} \frac{\partial x'^b}{\partial x'^d} \dot{e}_d \cdot \dot{e}_f \tag{18.106} \]
Now applying
\[ \dot{e}_d \cdot \frac{\partial \dot{e}_f}{\partial x'^g} = \Gamma^d_{fg}, \dot{e}_d \cdot \dot{e}_f = \delta^d_{g} \]
Eq. (18.106) becomes
\[ \Gamma'^b_{ac} = \frac{\partial x'^b}{\partial x'^d} \frac{\partial x^f}{\partial x'^a} \frac{\partial x^g}{\partial x'^c} \Gamma^d_{fg} + \frac{\partial x'^b}{\partial x'^d} \frac{\partial^2 x^d}{\partial x'^c \partial x'^a} \tag{18.107} \]
Homework: If we swap the derivatives with respect to $x$ and $x'$ in the expression

$$\frac{\partial x'^b}{\partial x^d} \frac{\partial^2 x'^d}{\partial x'^c \partial x'^a}$$

we will find

$$\frac{\partial x'^b}{\partial x^d} \frac{\partial^2 x'^d}{\partial x'^c \partial x'^a} = - \frac{\partial x^d}{\partial x'^a} \frac{\partial x'^f}{\partial x^d} \frac{\partial^2 x'^b}{\partial x'^c \partial x'^f}$$

show that we arrive at an alternative expression for the affine connect under coordinates transformations

$$\Gamma_{bc}^a = \frac{\partial x'^b}{\partial x^d} \frac{\partial x'^f}{\partial x^d} \frac{\partial x'^g}{\partial x'^c} \Gamma_{gd}^a - \frac{\partial x^d}{\partial x'^a} \frac{\partial x'^f}{\partial x^d} \frac{\partial^2 x'^b}{\partial x'^c \partial x'^f}$$ \hspace{1cm} (18.108)

The affine connection and the metric: we recall the affine connection

$$\Gamma_{ac}^b = \hat{e}^b \cdot \frac{\partial \hat{e}_a}{\partial x^c}$$ \hspace{1cm} (18.109)

in a similar manner we can also write

$$\Gamma_{ca}^b = \hat{e}^b \cdot \frac{\partial \hat{e}_c}{\partial x^a}$$ \hspace{1cm} (18.110)

The difference between Eq. (18.109) and (18.110), $T_{ac}^b$

$$T_{ac}^b = \Gamma_{ac}^b - \Gamma_{ca}^b$$ \hspace{1cm} (18.111)

is known as the torsion tensor. We will consider a torsionless manifolds, for which

$$T_{ac}^b = 0 \Rightarrow \Gamma_{ac}^b = \Gamma_{ca}^b.$$ \hspace{1cm} (18.112)

We will determine the relationship between the affine connection and the metric for a torsionless manifold. We recall the metric

$$g_{ab} = \hat{e}_a \cdot \hat{e}_b$$

so that

$$\frac{\partial g_{ab}}{\partial x^c} = \partial_c g_{ab} = \partial_c (\hat{e}_a \cdot \hat{e}_b) = \hat{e}_a \cdot \partial_c \hat{e}_b + \hat{e}_b \cdot \partial_c \hat{e}_a.$$ \hspace{1cm} (18.113)

Applying the relation

$$\partial_c \dot{e}_b = \Gamma^d_{bc} \dot{e}_d$$ \hspace{1cm} (18.114)

we can rewrite Eq. (18.113) as

$$\partial_c g_{ab} = \dot{e}_b \cdot \Gamma^d_{ac} \dot{e}_d + \dot{e}_a \cdot \Gamma^d_{bc} \dot{e}_d = \Gamma^d_{ac} \dot{e}_b \cdot \dot{e}_d + \Gamma^d_{bc} \dot{e}_a \cdot \dot{e}_d$$

$$\Rightarrow \partial_c g_{ab} = \Gamma^d_{ac} g_{bd} + \Gamma^d_{bc} g_{ad}.$$ \hspace{1cm} (18.115)
Similarly
\[ \partial_b g_{ca} = \partial_b (\hat{e}_c \cdot \hat{e}_a) = \hat{e}_c \cdot \partial_b \hat{e}_a + \hat{e}_a \cdot \partial_b \hat{e}_c \]
\[ \Rightarrow \partial_b g_{ca} = \hat{e}_c \cdot (\Gamma^d_{bc} \hat{e}_d + \hat{e}_a \cdot \Gamma^d_{ca} \hat{e}_d) \]
\[ \Rightarrow \partial_b g_{ca} = \Gamma^d_{ab} g_{cd} + \Gamma^d_{cb} g_{ad} \quad (18.116) \]
and
\[ \partial_a g_{bc} = \Gamma^d_{ba} g_{dc} + \Gamma^d_{ca} g_{bd} \quad (18.117) \]

Now combining Eqs. (18.115)-(18.117), we can write
\[ \partial_c g_{ab} + \partial_b g_{ca} - \partial_a g_{bc} = 2 \Gamma^d_{bc} g_{ad}. \quad (18.118) \]

Recalling that we will be interested in a torsionless manifold where \( \Gamma^b_{ac} = \Gamma^b_{ca} \), Eq. (18.118)
\[ \partial_c g_{ab} + \partial_b g_{ca} - \partial_a g_{bc} = 2 \Gamma^d_{bc} g_{ad}. \quad (18.119) \]

Multiplying Eq. (18.120) by \( g^{pa} \),
\[ g^{pa} (\partial_c g_{ab} + \partial_b g_{ca} - \partial_a g_{bc}) = 2 \Gamma^p_{bc} g^{pa} g_{ad} \quad (18.120) \]

From Eq. (18.64), we have
\[ g^{ab} g_{bc} = \delta^a_c \]
so that Eq. (18.64) becomes
\[ g^{pa} (\partial_c g_{ab} + \partial_b g_{ca} - \partial_a g_{bc}) = 2 \Gamma^p_{bc} g^{pa} \quad (18.121) \]

Since the summation is over \( d \), replacing \( d \) by \( p \), we find
\[ g^{pa} (\partial_c g_{ab} + \partial_b g_{ca} - \partial_a g_{bc}) = 2 \Gamma^p_{bc} \quad (18.122) \]

Now relabeling the index \( a \) by \( d \)
\[ g^{pd} (\partial_c g_{db} + \partial_b g_{cd} - \partial_d g_{bc}) = 2 \Gamma^p_{bc} \quad (18.123) \]
and then the index \( p \) by \( a \), we can write
\[ \Gamma^a_{bc} = \frac{g^{ad}}{2} (\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc}) \quad (18.124) \]

The right hand side of Eq. (18.124) is known as the **metric connection** and is often represented by \( \{_{bc}^a \}

**Useful Formulae:**
\[ \Gamma_{abc} = g_{ad} \Gamma^d_{bc} \quad (18.125) \]

Multiplying by \( g^{fa} \)
\[ g^{fa} \Gamma_{abc} = g^{fa} g_{ad} \Gamma^d_{bc} = \delta^f_d \Gamma^d_{bc} = \Gamma^f_{bc} \]
\[ \Rightarrow \Gamma^f_{bc} = g^{fa} \Gamma_{abc} \quad (18.126) \]
Applying the relation in Eq (18.124), we can express Eq. (18.125) as
\[
\Gamma^f_{bc} = \frac{g^{fd}}{2} (\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc}) \tag{18.127}
\]
Relabeling the index \(d\) by \(a\) in Eq. (18.127), we have
\[
\Gamma^f_{bc} = \frac{g^{fa}}{2} (\partial_b g_{ca} + \partial_c g_{ab} - \partial_a g_{bc}) \tag{18.128}
\]
so that equating Eqs. (18.126) and (18.128), we find
\[
\Gamma_{abc} = \frac{1}{2} (\partial_b g_{ca} + \partial_c g_{ab} - \partial_a g_{bc}) \tag{18.129}
\]
The quantity \(\Gamma_{abc}\) is traditionally known as a Christoffel symbol of the first kind. Noting that
\[
\Gamma_{bac} = \frac{1}{2} (\partial_a g_{cb} + \partial_c g_{ba} - \partial_b g_{ac}) \tag{18.130}
\]
we can write
\[
\Gamma_{abc} + \Gamma_{bac} = \frac{1}{2} (\partial_b g_{ca} + \partial_c g_{ab} - \partial_a g_{bc} + \partial_a g_{cb} + \partial_c g_{ba} - \partial_b g_{ac}) \tag{18.131}
\]
taking into account the symmetry of the metric tensor, from Eq. (18.131), we find
\[
\partial_c g_{ab} = \Gamma_{abc} + \Gamma_{bac} \tag{18.132}
\]
Eq. (18.132) allows us to express the partial derivative of the metric components in terms of the connection coefficients.

We recall that the determinant of the matrix \(G\) can be expressed as
\[
\det G = g_{ab} (-1)^{a+b} M_{ab} = g_{ab} \text{cof} [G]_{ab}, \tag{18.133}
\]
where
\[
[\text{cof} [G]]_{ab} = (-1)^{a+b} M_{ab} \tag{18.134}
\]
is the cofactor matrix to \(G\) which is determined from the minor, \(M_{ab}\) and it is a constant matrix. We recall that the minor of matrix \(G\) denoted by \(M_{ab}\) is the determinant of the matrix formed from matrix \(G\) by removing the \(a^{th}\) row and \(b^{th}\) column. Let \(\det G = g\), so that
\[
g = g_{ab} \text{cof} [G]_{ab} \Rightarrow \partial_c g = \text{cof} [G]_{ab} \partial_c g_{ab}. \tag{18.135}
\]
Noting that the cofactor matrix can be expressed as
\[
g^{ab} g = g^{ab} g_{ab} \text{cof} [G]_{ab} \Rightarrow \text{cof} [G]_{ab} = g^{ab} g \tag{18.136}
\]
we may write
\[
\partial_c g = \text{cof} [G]_{ab} \partial_c g_{ab} = g g^{ab} \partial_c g_{ab}. \tag{18.137}
\]
Now using the relation (18.132), we can write
\[ \partial_c g = gg^{ab} \partial_c g_{ab} = gg^{ab} (\Gamma_{abc} + \Gamma_{bac}). \] (18.138)

We recall
\[ \Gamma_{abc} = g_{ad} \Gamma_{d bc} \] (18.139)

so that
\[ \partial_c g = gg^{ab} (\Gamma_{abc} + \Gamma_{bac}) = g (g^{ab} g_{ad} \Gamma_{d bc} + g^{ab} g_{bd} \Gamma_{d ac}) = g \left( \delta^b_d \Gamma_{d bc} + \delta^a_d \Gamma_{d ac} \right). \] (18.140)

Taking into account that \(a\) and \(b\) are dummy indices, we can replace \(b\) by \(a\) so that
\[ \partial_c g = 2g \Gamma_{ac} \] (18.141)

This can be rearranged as
\[ \Gamma_{ac} = \frac{1}{2g} \partial_c g = \frac{1}{2} \partial_c \ln |g| = \partial_c \ln \sqrt{|g|}. \] (18.142)

Now for the sake of convenience if we replace \(c\) by \(b\), we may write the above equation as
\[ \Gamma_{ab} = \partial_b \ln \sqrt{|g|}. \] (18.143)

The modulus is for the case where the manifold is seudo-Riemannian where the metric elements can be negative.

### 18.6 Local geodesic and Cartesian coordinates

Let’s consider a manifold with coordinates system, \(x^a\), and a point \(P\) on this manifold with coordinates, \(x^a_P\). Let’s now define a new system of coordinates, \(x^a_0\), in terms of \(x^a_P\), and the coordinates \(x^a\) as
\[ x^a_0 = x^a - x^a_P + \frac{1}{2} \Gamma_{bc}^a (P) (x^b - x^b_P) (x^c - x^c_P). \] (18.144)

We must know that under this transformation how the affine connection transformed. We recall that under coordinate transformation \(x^a \rightarrow x'^a\), the affine connection is transformed according to
\[ \Gamma'^{bf}_{ac} = \frac{\partial x'^b}{\partial x^a} \frac{\partial x'^f}{\partial x^e} \Gamma_{d e c}^{bd} - \frac{\partial x'^d}{\partial x^a} \frac{\partial x'^f}{\partial x^e} \frac{\partial x'^b}{\partial x^d} \frac{\partial x'^{bf}}{\partial x^f}. \] (18.145)

In order to determine this at point \(P\), we need to differentiate Eq. (18.144)
\[ \frac{\partial x'^a}{\partial x^d} = \frac{\partial x^a}{\partial x^d} + \frac{1}{2} \Gamma_{bc}^a (P) \frac{\partial}{\partial x^d} \left( (x^b - x^b_P) (x^c - x^c_P) \right) \]
\[ = \frac{\partial x^a}{\partial x^d} + \frac{1}{2} \Gamma_{bc}^a (P) \left\{ (x^c - x^c_P) \frac{\partial x^b}{\partial x^d} + (x^b - x^b_P) \frac{\partial x^c}{\partial x^d} \right\}. \] (18.146)
CHAPTER 18. VECTOR CALCULUS ON MANIFOLDS

where we used the fact that \( \Gamma^a_{bc} (P) \), \( x^a_P \), \( x^b_P \), and \( x^c_P \) are constant at point \( P \). The coordinates \( x^a \) are independent coordinates,

\[
\frac{\partial x^a}{\partial x^d} = \delta^a_d, \quad \frac{\partial x^b}{\partial x^d} = \delta^b_d, \quad \frac{\partial x^c}{\partial x^d} = \delta^c_d
\]  

(18.147)

so that

\[
\frac{\partial x'^a}{\partial x'^d} = \delta^a_d + \frac{1}{2} \Gamma^a_{bc} (P) \left\{ (x'^c - x^c_P) \delta^b_d + (x^b - x^b_P) \delta^c_d \right\}
\]

(18.148)

which results in

\[
\frac{\partial x'^a}{\partial x'^d} = \delta^a_d + \frac{1}{2} \Gamma^a_{dc} (P) (x'^c - x^c_P) + \Gamma^a_{bd} (P) (x^b - x^b_P)
\]

(18.149)

The summation indices are dummy indices and therefore we can replace \( b \) by \( c \) in the second term so that

\[
\frac{\partial x'^a}{\partial x'^d} = \delta^a_d + \frac{1}{2} \left[ \Gamma^a_{dc} (P) (x'^c - x^c_P) + \Gamma^a_{cd} (P) (x^c - x^c_P) \right]
\]

(18.150)

Since for a torsionless manifold

\[ \Gamma^a_{dc} (P) = \Gamma^a_{cd} (P), \]

we can write

\[
\frac{\partial x'^a}{\partial x'^d} = \delta^a_d + \Gamma^a_{dc} (P) (x'^c - x^c_P)
\]

(18.151)

We are interested in the affine connection at point \( P \) so that when we are evaluating Eq. (18.151) at \( x^c = x^c_P \), we find

\[
\left. \frac{\partial x'^a}{\partial x'^d} \right|_P = \delta^a_d
\]

(18.152)

Similarly, the inverse is also given by

\[
\left. \frac{\partial x^a}{\partial x^d} \right|_P = \delta^a_d
\]

(18.153)

Differentiating Eq. (18.151) with respect to \( x^e \), we have

\[
\frac{\partial x'^a}{\partial x'^d} \frac{\partial x^e}{\partial x^d} = \frac{\partial}{\partial x^e} \delta^a_d + \Gamma^a_{dc} (P) \frac{\partial}{\partial x^e} (x'^c - x^c_P)
\]

\[
\Rightarrow \quad \frac{\partial x'^a}{\partial x'^d} \frac{\partial x^e}{\partial x^d} = \Gamma^a_{dc} (P) \delta^e_c = \Gamma^a_{dc} (P)
\]

(18.154)

Using the results in Eqs. (18.152)-(18.154), the transformation equation for the affine connection

\[
\Gamma^b_{ac} = \frac{\partial x'^b}{\partial x^d} \frac{\partial x^f}{\partial x^a} \frac{\partial x^g}{\partial x^c} \Gamma^a_{df} - \frac{\partial x^d}{\partial x^a} \frac{\partial x^f}{\partial x^c} \frac{\partial^2 x'^b}{\partial x^d \partial x^f}
\]

(18.155)
at point $P$ becomes
\[ \Gamma^{ab}_{ac} (P) = \delta^b_d \delta^f_e \delta^g_f \Gamma^d_{fg} (P) - \delta^d_e \delta^f_c \Gamma^b_{df} (P) \] (18.156)
which simplified into
\[ \Gamma^{ab}_{ac} (P) = \delta^b_d \delta^f_e \Gamma^d_{fc} (P) - \delta^d_e \Gamma^b_{dc} (P) = \delta^b_d \Gamma^d_{ac} (P) - \Gamma^b_{ac} (P) \] (18.157)
\[ \Rightarrow \Gamma^{ab}_{ac} (P) = \Gamma^b_{ac} (P) - \Gamma^b_{ac} (P) = 0. \]

The result in Eq. (18.157) shows that for the coordinate transformation defined by Eq. (18.144), the affine connection becomes zero. Such coordinates where the affine connection becomes zero at a point $P$ on a manifold is known as \textit{local geodesic coordinates} about $P$.

In chapter 2 we have shown that the conditions for local Cartesian coordinates at a given point $P$ in a seudo-Riemmanian manifold are
\[ g'_{ab} (P) = \eta_{ab}, \] (18.158)
\[ \left. \frac{\partial g'_{ab} (x')}{\partial x'^c} \right|_P = 0 \] (18.159)
where $[\eta_{ab}] = diag (\pm 1, \pm 1, \ldots, \pm 1)$. We have also learned that the number of positive entries $(N_+)$ minus the number of negative entries $(N_-)$ in $[\eta_{ab}]$ is called the \textit{signature} of the manifold. For geodesic coordinates Eq. (18.159) can easily be shown applying the equation that relates the connection with the metric in Eq. (18.115). For the $x'^a$ coordinates, Eq. (18.115) can be written as
\[ \partial_c g'_{ab} = \Gamma^{cf}_{ac} g'_{bd} + \Gamma^{cf}_{bc} g'_{ad} \] (18.160)

When this equation is evaluated at point $P$
\[ \partial_c g'_{ab} = \left. \frac{\partial g'_{ab} (x')}{\partial x'^c} \right|_P = \Gamma^{cf}_{ac} (P) g'_{bd} (P) + \Gamma^{cf}_{bc} (P) g'_{ad} (P) \] (18.161)
and for a geodesic coordinates the connection is zero at point $P$, and therefore
\[ \left. \frac{\partial g'_{ab} (x')}{\partial x'^c} \right|_P = 0. \]

It is important to note that for geodesic coordinates the metric does not necessarily satisfy Eq. (18.158). But we can find coordinates $x'^{ra}$ that satisfy Eq. (18.158) by making a linear transformation to the $x'^{ra}$
\[ x'^{ra} = X x^{ra} \] (18.162)
where $X^a_b$ are constants.
18.7 The gradient, the divergence, the curl on a manifold

Before we see how the gradient of a scalar, the divergence or the curl of a vector is determined on a manifold, first we need to know how a vector is differentiated. Consider a vector, $\vec{v}$, in terms of its contravariant components

$$\vec{v} = v^a \hat{e}_a,$$  \hspace{1cm} (18.163)

where $\hat{e}_a$ are the coordinate basis vectors. The derivative of this vector with respect to the coordinate, $x^b$, can be expressed as

$$\frac{\partial \vec{v}}{\partial x^b} = \partial_b \vec{v} = \partial_b (v^a \hat{e}_a) = \hat{e}_a \partial_b (v^a) + v^a \partial_b (\hat{e}_a)$$  \hspace{1cm} (18.164)

We recall

$$\partial_b (\hat{e}_a) = \Gamma^c_{ab} \hat{e}_c$$

so that

$$\partial_b \vec{v} = (\partial_b v^a) \hat{e}_a + \Gamma^c_{ab} v^a \hat{e}_c.$$ 

(18.165)

Switching the places for the indices $c$ and $a$ one can write

$$\Gamma^c_{ab} v^a \hat{e}_c = \Gamma^a_{cb} v^c \hat{e}_a$$ 

(18.166)

so that

$$\partial_b \vec{v} = (\partial_b v^a) \hat{e}_a + v^c \Gamma^a_{cb} \hat{e}_a = (\partial_b v^a + \Gamma^a_{cb} v^c) \hat{e}_a.$$ 

(18.167)

The quantity in the bracket which is represented as

$$\nabla_b v^a = \partial_b v^a + \Gamma^a_{cb} v^c$$ 

(18.168)

is known as the covariant derivative of the vector components. Thus the derivative of a vector can be expressed as

$$\partial_b \vec{v} = (\nabla_b v^a) \hat{e}_a$$ 

(18.169)

For geodesic coordinates where the affine connection vanishes,

$$\Gamma^a_{cb} = 0$$

the covariant derivative reduces to

$$\nabla_b v^a = \partial_b v^a$$ 

(18.170)

which is just the ordinary derivative that we are very familiar with!

**Homework:** Suppose the vector is expressed in terms of its covariant components

$$\vec{v} = v^a \hat{e}_a$$ 

(18.171)

show that

$$\nabla_b v^a = \partial_b v^a - \Gamma^a_{cb} v^c$$ 

(18.172)
The covariant derivative of a scalar function: for a scalar function $\phi$ the covariant derivative is
\[ \nabla_b \phi = \partial_b \phi \] (18.173)

The gradient: The gradient of a scalar function $\phi$ is given by
\[ \nabla \phi = (\partial_a \phi) \hat{e}^a \] (18.174)

The divergence: The divergence of a vector field expressed in terms of its covariant components
\[ \vec{v} = v^a \hat{e}_a \] (18.175)
is given by
\[ \nabla \cdot \vec{v} = \nabla_a v^a \] (18.176)
Using the relation we obtained
\[ \nabla_b v^a = \partial_b v^a + \Gamma^a_{bc} v^c \] (18.177)
for $b = a$ and replacing $c$ by $b$, we find
\[ \nabla \cdot \vec{v} = \nabla_a v^a = \partial_a v^a + \Gamma^a_{ba} v^b. \] (18.178)
Using the relation
\[ \Gamma^a_{ab} = \partial_b \ln \sqrt{|g|} = \frac{1}{\sqrt{|g|}} \partial_a \sqrt{|g|} \] (18.179)
we can write
\[ \nabla \cdot \vec{v} = \partial_a v^a + v^a \frac{1}{\sqrt{|g|}} \partial_a \sqrt{|g|} = \frac{1}{\sqrt{|g|}} \partial_a \left[ v^a \sqrt{|g|} \right] \] (18.180)

The Laplacian: We recall that in the Euclidean space the Laplacian of the scalar function $\phi$, is given by
\[ \nabla^2 \phi = \nabla \cdot \nabla \phi. \] (18.181)
Applying Eq. (18.174) we can write
\[ \nabla^2 \phi = \nabla \cdot [(\partial_a \phi) \hat{e}^a]. \] (18.182)
But in order to apply the relation we derive for the divergence in Eq. (18.180), we need the vector
\[ \vec{v} = (\partial_a \phi) \hat{e}^a = v_a \hat{e}^a. \] (18.183)
in terms of its covariant components. We have seen that the index can be raised or lowered using the metric tensor. In this case we want to raise it, so we have
\[ g^{ab} v_b = v^a \]
and we can express the vector $\vec{v}$ as
\[ \vec{v} = v^a \hat{e}_a = g^{ab} (\partial_b \phi) \hat{e}_a. \] (18.184)
CHAPTER 18. VECTOR CALCULUS ON MANIFOLDS

Then the Laplacian becomes
\[ \nabla^2 \phi = \nabla \cdot \left[ g^{ab} \left( \partial_b \phi \right) \hat{e}_a \right]. \quad (18.185) \]

Now applying the relation
\[ \nabla \cdot \vec{v} = \frac{1}{\sqrt{|g|}} \partial_a \left[ v^a \sqrt{|g|} \right] \quad (18.186) \]

we may write the Laplacian
\[ \nabla^2 \phi = \nabla_a \nabla^a \phi = \frac{1}{\sqrt{|g|}} \partial_a \left[ \sqrt{|g|} g^{ab} \partial_b \phi \right]. \quad (18.187) \]

The Laplacian symbol \( \nabla^2 \) is used in the usual 3-D Euclidean space or in an N-D manifold. In 4-D spacetime manifold, as you will see (or have seen) in the relativistic electrodynamics in *Theoretical Physics IV (Electricity & Magnetism II)*, \( \nabla^2 \) is replaced by \( \Box^2 \) known as the *d’Alembertian operator*.

**Curl:** The curl is defined as a rank-2 antisymmetric tensor with components
\[ (\text{curl} \vec{v})_{ab} = \nabla_a v_b - \nabla_b v_a. \quad (18.188) \]

Using the relation
\[ \nabla_b v_a = \partial_b v_a - \Gamma^c_{ab} v^c \quad (18.189) \]

we can express the curl as
\[ (\text{curl} \vec{v})_{ab} = \partial_a v_b - \Gamma^c_{ab} v^c - \partial_b v_a + \Gamma^c_{ab} v^c = \partial_a v_b - \partial_b v_a. \quad (18.190) \]

**18.8 Intrinsic derivative of a vector along a curve**

We will encounter vector fields that does depend on a curve instead of the entire or some region of the manifold. In such cases the curve may be defined by the coordinates \( x^a \) that depends on some parameter, \( u \). Let’s consider a vector, \( v \), expressed in terms of its controvariant components. Since the coordinates on this curve depends on the parameter, \( u \), this vector can be expressed in terms of this parameter as
\[ \vec{v}(u) = v^a(u) \hat{e}_a(u). \quad (18.191) \]

The derivative of this vector along this curve is given by
\[
\frac{d}{du} \vec{v}(u) = \frac{d}{du} \left[ v^a(u) \hat{e}_a(u) \right] = v^a(u) \frac{d\hat{e}_a(u)}{du} + \hat{e}_a(u) \frac{dv^a(u)}{du}.
\]

\[
\Rightarrow \frac{d}{du} \vec{v}(u) = v^a(u) \frac{d\hat{e}_a(u)}{dx^b} \frac{dx^b}{du} + \hat{e}_a(u) \frac{dv^a(u)}{du}.
\]

Using the relation
\[ \frac{d\hat{e}_a}{dx^b} = \Gamma^d_{ab} \hat{e}_d \quad (18.194) \]
18.9. PARALLEL TRANSPORT

we have

$$\frac{d}{du} \vec{v} = \nu^a \frac{dx^b}{du} \Gamma^d_{ab} \dot{e}_d + \dot{e}_a \frac{dv^a}{du}$$ (18.195)

and replacing the index $a$ by $c$ in the first term

$$\frac{d}{du} \vec{v} = \nu^c \frac{dx^b}{du} \Gamma^d_{cb} \dot{e}_d + \dot{e}_a \frac{dv^a}{du}$$ (18.196)

and $d$ by $a$

$$\frac{d}{du} \vec{v}(u) = \nu^c \frac{dx^b}{du} \Gamma_{cb}^a \dot{e}_a + \dot{e}_a \frac{dv^a}{du} = \left( \frac{dv^a}{du} + \Gamma_{cb}^a \frac{dx^b}{du} \right) \dot{e}_a,$$ (18.197)

which we put in the form

$$\frac{d}{du} \vec{v}(u) = \frac{Dv^a}{Du} \dot{e}_a,$$ (18.198)

where

$$\frac{Dv^a}{Du} = \frac{dv^a}{du} + \Gamma_{cb}^a \frac{dx^b}{du}$$ (18.199)

is called the intrinsic (or absolute) derivative of the component $v^a$. Substituting

$$\frac{dv^a}{du} = \frac{dv^a}{dx^b} \frac{dx^b}{du}$$ (18.200)

into Eq. (18.199), we find

$$\frac{Dv^a}{Du} = \frac{\partial v^a}{\partial x^b} \frac{dx^b}{du} + \Gamma_{cb}^a \frac{dx^b}{du} = \frac{\partial v^a}{\partial x^b} \frac{dx^b}{du} + \Gamma_{cb}^a \frac{dx^b}{du} = \left( \frac{\partial v^a}{\partial x^b} + \Gamma_{cb}^a \frac{dx^b}{du} \right) \frac{dx^b}{du}$$

$$\Rightarrow \frac{Dv^a}{Du} = \left( \nabla_b v^a \right) \frac{dx^b}{du},$$ (18.201)

where we used the relation in Eq. (18.168).

Homework: Suppose the vector, $\vec{v}$, depends on the parameter $u$ on a curve defined by $x^a(u)$ is expressed in terms of its covariant components

$$\vec{v} = v_a(u) \dot{e}_a(u).$$ (18.202)

Show that the intrinsic derivative of this vector is given by

$$\frac{Dv_a}{Du} = \frac{dv_a}{du} - \Gamma_{ac}^b \frac{dx^c}{du}$$ (18.203)

18.9 Parallel transport

In order to understand the idea of parallel transport of a vector on a manifold let’s consider motion of a particle in space. Suppose the position of the particle depends on time, $t$, then the displacement is parametrized by time $t$, $D(t)$. The velocity of the particle is given by

$$\vec{v} = \frac{dD}{dt}$$ (18.204)
and the acceleration by

\[ \ddot{a} = \frac{d\dot{v}}{dt} \]  

(18.205)

Suppose you plot the displacement of the particle at different times, then you would get generally a curve. The particle would have a constant velocity through this curve provided its acceleration is zero.

\[ \ddot{a} = \frac{d\dot{v}}{dt} = 0. \]  

(18.206)

This means the particle travels along this curve with a constant velocity. The velocity would have the same magnitude and direction. The velocity vector remains parallel at each point on the curve describing the displacement of the particle as a function of time.

On a curve \( C \) on a manifold, a parallel transport of a vector \( \dot{v} = v_a (u) \dot{e}^a (u) \) (18.207) is when the intrinsic derivative of this vector is zero

\[ \frac{Dv_a}{Du} = \frac{dv_a}{du} - \Gamma^b_{ac} v_b \frac{dx^c}{du} = 0 \]  

(18.208)

### 18.10 Null curves, non-null curves, and affine parameter

We recall that the tangent vector \( t \) at point \( p \) on a manifold is the vector that lies in the tangent space \( T_p \) at that point and is given by

\[ \dot{t} = \lim_{\delta u \to 0} \frac{\partial \vec{s}}{\partial u} = \frac{d\vec{s}}{du}, \]  

(18.209)

where \( \partial \vec{s} \) is the infinitesimal separation vector between the point \( P \) and some nearby point \( Q \) on the curve on the manifold corresponding to the parameter value \( u + \delta u \).

In a given coordinate system \( x^a \) with basis vectors \( \dot{e}_a \), we can express the infinitesimal separation vector \( d\vec{s} \) as

\[ d\vec{s} = \frac{dx^a}{du} \dot{e}_a \]  

(18.210)

so that the tangent vector becomes

\[ \dot{t} = \frac{dx^a}{du} \dot{e}_a \]  

(18.211)

Recalling that in pseudo-Riemannian manifold the length of a vector \( \dot{v} \) is given by

\[ |\dot{v}| = \sqrt{|g_{ab}v^av^b|} = \sqrt{|g^{ab}v_av_b|} = \sqrt{|v^av_b|} = \sqrt{|v_a v^b|} \]  

(18.212)
null curves, non-null curves, and affine parameter

18.10. Figure 18.5: A tangent vector at point $P$ in the tangent space.

we have

$$|t| = \sqrt{g_{ab}t^a t^b} = \sqrt{g^{ab} \frac{dx^a}{du} \frac{dx^b}{du}} = \frac{\sqrt{|g^{ab}dx^a dx^b|}}{du}$$

we recall the metric or the interval (the distance squared along the curve on the manifold between the two points $P$ and $Q$) is

$$ds^2 = g^{ab}dx^a dx^b$$

Eq. (18.213) becomes

$$|t| = \left| \frac{ds}{du} \right|$$

The non-null vectors and the affine parameter: for the non-null vectors $|t| \neq 0$. This means according to Eq. (18.214) the distance $ds$ at all points on the curve must be different from zero and therefore it depends on the parameter $u$ at all points on the curve, $s = s(u)$. If parameter $u$ and the distance $s$ are related by

$$u = as + b$$

for $a, b \neq 0$, the parameter $u$ is called the affine parameter on the curve.

The null vectors: if the tangent vector is a null vector, then

$$|t| = \left| \frac{ds}{du} \right| = 0$$

at all points on the curve and the distance, $s$, does not depend on the parameter, $u$, and we clearly can not use it as affine parameter since it does not satisfy the condition in Eq. (18.215). But it is possible to find a privileged family of affine parameter.
18.11 Refreshment from theoretical physics: the calculus of variation

Geodesic: The curve along a surface which marks the shortest distance between two neighboring points. Finding geodesics is one of the problems which we can solve using the calculus of variation.

Stationary point: A point on a given function \( f(x) \) is said to be stationary point when
\[
\frac{df(x)}{dx} = 0.
\]

Geodesic on an Euclidean 2-D space: Consider two points in a x-y plane \( P_1 \) and \( P_2 \). Prove that the shortest distance between the two points is the distance measured along a straight line.

Let’s consider two points on the x-y plane. Let first point be \((x_1, y_2)\) and the second point be \((x_1, y_2)\). Then the distance between these points is given by the integral
\[
I = \int_{(1)}^{(2)} ds
\]
where
\[
ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \sqrt{1 + \left(\frac{dx}{dy}\right)^2} dy
\]
(18.218)

We may rewrite this distance as
\[
I = \int_{(1)}^{(2)} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx.
\]
(18.219)

Out of the infinitely many functions that can be used to connect the two points, we want to determine the one that would give the minimum distance. Let these function be denoted by \( Y(x) \). From these infinite number of functions there is only one function that gives the minimum distance between the two points. If this function is \( y(x) \), then we may write \( Y(x) \) in terms of \( y(x) \) as
\[
Y(x) = y(x) + \epsilon \eta(x)
\]
(18.220)
where \( \eta(x) \) is an arbitrary function which must satisfy the condition
\[
\eta(x_1) = \eta(x_2) = 0
\]
(18.221)
so that
\[
Y(x) = y(x)
\]
(18.222)
at the two points \((x = x_1 = x_2)\); \( \epsilon \) is the constant of variation that means it is constant that determines by how much \( Y(x) \) differs from \( y(x) \). Now in terms of \( Y \), we may write
\[
I(\epsilon) = \int_{(1)}^{(2)} \sqrt{1 + Y'^2} dx.
\]
(18.223)
where

\[ Y' = \frac{dY(x)}{dx}. \]  

(18.224)

For the distance to be a minimum, the necessary condition is that the integral must be stationary. This requires

\[ \frac{dI(\epsilon)}{d\epsilon} \bigg|_{\epsilon=0} = 0 \]  

(18.225)

which leads to

\[ \frac{dI(\epsilon)}{d\epsilon} \bigg|_{\epsilon=0} = \int_{(1)}^{(2)} \left( -\frac{1}{2} \right) \frac{1}{\sqrt{1 + Y'^2}} \left( \frac{dY''(\epsilon)}{d\epsilon} \right) dx \bigg|_{\epsilon=0} = 0. \]  

Using

\[ Y(x) = y(x) + \epsilon \eta(x) \]  

(18.226)

we may write

\[ Y'(x) = y'(x) + \epsilon \eta'(x) \Rightarrow Y'(x) \bigg|_{\epsilon=0} = y'(x) \]  

(18.227)

so that

\[ \frac{dY''(\epsilon)}{d\epsilon} = \eta'(x) \Rightarrow \frac{dY''(\epsilon)}{d\epsilon} \bigg|_{\epsilon=0} = \eta'(x) \]  

(18.228)

which leads to

\[ \int_{(1)}^{(2)} \left( -\frac{1}{2} \right) \frac{1}{\sqrt{1 + Y'^2}} \left( \frac{dY''(\epsilon)}{d\epsilon} \right) dx \bigg|_{\epsilon=0} = \int_{(1)}^{(2)} \frac{y' \eta'(x)}{\sqrt{1 + y'^2}} dx = 0. \]  

(18.229)

Using integration by parts

\[ \int u dv = uv - \int v du \]  

(18.230)

for

\[ \eta'(x) = dv \Rightarrow v = \eta(x), u = \frac{y'}{\sqrt{1 + y'^2}} \Rightarrow du = \frac{d}{dx} \left( \frac{y'}{\sqrt{1 + y'^2}} \right) dx \]  

(18.231)

we may write Eq. (18.229) as

\[ \int_{(1)}^{(2)} \frac{y' \eta'(x)}{\sqrt{1 + y'^2}} dx = \frac{y'}{\sqrt{1 + y'^2}} \eta(x) \bigg|_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta(x) \frac{d}{dx} \left( \frac{y'}{\sqrt{1 + y'^2}} \right) dx = 0. \]  

(18.232)

Using the condition

\[ \eta(x_1) = \eta(x_2) = 0 \]
we find
\[ \frac{dI(e)}{de} \bigg|_{e=0} = \int_{x_1}^{x_2} \frac{y'y''(x)}{\sqrt{1+y'^2}} \, dx = - \int_{x_1}^{x_2} \eta(x) \frac{dy'}{dx} \left( \frac{y'}{\sqrt{1+y'^2}} \right) \, dx = 0 \]
\[ (18.233) \]

Since \( \eta(x) \) is an arbitrary function, we must have
\[ \frac{dy'}{dx} \left( \frac{y'}{\sqrt{1+y'^2}} \right) = 0 \]
\[ \Rightarrow \frac{dy'}{dx} (y'^2) = 0 \Rightarrow \frac{dy'}{dx} = 0 \Rightarrow y' = m \]  
\[ (18.234) \]
where \( m \) is a constant. Thus
\[ y' = \frac{dy}{dx} = m \Rightarrow y(x) = mx + b \]
which is equation of a straight line.

The Problem: We consider some unknown function \( y(x) \). We assume that this function is known at two fixed points, \( y(x_1) \) and \( y(x_2) \). We wish to find the function \( y(x) \) that makes the integral
\[ I = \int_{x_1}^{x_2} F(x, y, y') \, dx \]
stationary for some known function \( F(x, y, y') \). It can be shown that this function satisfies the Euler-Lagrange Equation:
\[ \frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right) - \frac{\partial F}{\partial y} = 0 \]

18.12 The geodesic

We recall that a local geodesic at point \( P \) is where the affine connection is zero
\[ \Gamma^b_{ac} (P) = \Gamma^b_{ac} (P) - \Gamma^b_{ac} (P) = 0. \]  
\[ (18.236) \]
Suppose we have a set of points in Euclidean space defining a geodesic curve, then at all of these points the affine connection must vanish.
\[ \partial_c \hat{e}_b = \Gamma^f_{bc} \hat{e}_f = 0 \]
\[ (18.237) \]
The basis vector with respect to the coordinates \( x^c \) does not change along the geodesic curve. This means for the tangent vector
\[ \hat{t} = \frac{dx^a}{d\lambda} \hat{e}_a \]
\[ (18.238) \]
to the curve at least the direction remains the same. In the Euclidean space this makes the curve to be a straight line where the tangent vectors has same direction along the line. Thus for Euclidean space the geodesic is a straight line.

For general curve defined by $x^a = x^a(u)$, on a manifold, if the curve is geodesic then the tangent curve must have the same direction at all points on the curve. This mean the change in the tangent vector with respect to the parameter $u$ is then only the magnitude that changes and it is given by

$$\frac{d\hat{t}}{du} = \lambda(u) \hat{t} = \lambda(u) \frac{dx^a}{du} \hat{e}_a, \quad (18.239)$$

where $\lambda(u)$ is some function of $u$. Using our result for intrinsic derivative of a vector $\vec{v}(u)$

$$\frac{d}{du} \vec{v}(u) = \frac{Dv^a}{Du} \hat{e}_a = \left( \frac{dv^a}{du} + \Gamma^a_{cb} v^c \frac{dx^b}{du} \right) \hat{e}_a, \quad (18.240)$$

for the tangent vector intrinsic derivative, we have

$$\frac{d\hat{t}}{du} = \frac{Dt^a}{Du} \hat{e}_a = \left( \frac{dt^a}{du} + \Gamma^a_{cb} t^c \frac{dx^b}{du} \right) \hat{e}_a, \quad (18.241)$$

so that substituting this into Eq. (18.239), we find

$$\frac{dt^a}{du} + \Gamma^a_{cb} t^c \frac{dx^b}{du} = \lambda(u) \frac{dx^a}{du} \quad (18.242)$$

Noting that

$$\hat{t} = \frac{dx^a}{du} \hat{e}_a = t^a \hat{e}_a \quad (18.243)$$

Eq. (18.242) can be written as

$$\frac{d^2 x^a}{du^2} + \Gamma^a_{cb} \frac{dx^c}{du} \frac{dx^b}{du} = \lambda(u) \frac{dx^a}{du} \quad (18.244)$$

The result in Eq. (18.244) is valid for both non-null and null geodesics parameterized in terms of some general parameter $u$. For affine parameter $u$ where it is related to the distance $s$ on the curve

$$u = as + b \quad (18.245)$$

we have

$$du = ads \quad (18.246)$$

and

$$|\hat{t}| = \left| \frac{ds}{du} \right| = a \quad (18.247)$$

which is a tangent vector with a constant length that is independent of the parameter $u$. This means

$$\frac{d\hat{t}}{du} = \lambda(u) \hat{t} = 0 \Rightarrow \lambda(u) = 0 \quad (18.248)$$
Therefore in general for a privilege parameter called the affine parameter the equation for the geodesic in Eq. (18.244) can be written as

\[
\frac{d^2x^a}{du^2} + \Gamma^a_{cb} \frac{dx^c}{du} \frac{dx^b}{du} = 0
\]  

(18.249)

Eq. (18.249) is a parallel transport for the tangent vector that we discussed in the previous section.

\[
\frac{Dx^a}{Du} = \frac{dt^a}{du} + \Gamma^a_{cb} \frac{dx^c}{du} \frac{dx^b}{du} = 0
\]  

(18.250)

which can be shown by replacing

\[
\frac{dx^a}{du} = t^a, \quad \frac{dx^c}{du} = t^c
\]  

(18.251)

in Eq. (18.249).

**Homework:**

If we change the affine parameter \(u\) to \(ut\), the coordinates that define the geodesic curve would change from \(x^a(u)\) to \(x^a(u')\). Show that in terms of the new affine parameter, \(ut\), the geodesic in Eq. (18.249) becomes

\[
\frac{d^2x^a}{du'^2} + \Gamma^a_{cb} \frac{dx^c}{du'} \frac{dx^b}{du'} = \left(\frac{d^2u}{du^2}\right) \frac{dx^a}{du'}
\]  

(18.252)

### 18.13 Stationary property of the non-null geodesic

Let’s consider the curve \(C\) in our manifold defined by the coordinates \(x^a(u)\). Suppose we have two points 1 and 2 on this curve and we are interested in the length along this curve joining these two points, this length can be determined from

\[
L = \int_1^2 ds = \int_1^2 \sqrt{|g_{ab}dx^a dx^b|} = \int_1^2 \sqrt{g_{ab} \frac{dx^a}{du} \frac{dx^b}{du}} du
\]  

(18.253)

or using the notation

\[
\frac{dx^a}{du} = \dot{x}^a
\]  

(18.254)

this length can be expressed as

\[
L = \int_1^2 \sqrt{|g_{ab} \dot{x}^a \dot{x}^b|} du = \int_1^2 F du,
\]  

(18.255)

where

\[
F = \sqrt{|g_{ab} \dot{x}^a \dot{x}^b|} = \dot{s} = \frac{ds}{du}
\]  

(18.256)
Using the principle of variation that you have learned in *Theoretical Physics Part I* and revised in the previous section the curve along a surface which marks the shortest distance between two neighboring points is the geodesic. Using the principle of variation it can be shown that the integral must be stationary for the geodesic and the function satisfies the Euler-Lagrange Equation:

\[
\frac{d}{du} \left( \frac{\partial F}{\partial \dot{x}^c} \right) - \frac{\partial F}{\partial x^c} = 0.
\]

(18.257)

Using Eq. (18.256), we have

\[
\frac{\partial F}{\partial x^c} = \frac{\partial}{\partial x^c} \left[ \sqrt{|g_{ab}\dot{x}^a\dot{x}^b|} \right] = \frac{\dot{x}^a\dot{x}^b \partial g_{ab}}{2\sqrt{|g_{ab}\dot{x}^a\dot{x}^b|}} = \frac{\dot{x}^a\dot{x}^b \partial g_{ab}}{2\hat{s}}
\]

\[
\frac{\partial F}{\partial \dot{x}^c} = \frac{\partial}{\partial \dot{x}^c} \left[ \sqrt{|g_{ab}\dot{x}^a\dot{x}^b|} \right] = \frac{g_{ab}}{2\sqrt{|g_{ab}\dot{x}^a\dot{x}^b|}} \left\{ \dot{x}^a \frac{\partial \dot{x}^a}{\partial \dot{x}^c} + \dot{x}^a \frac{\partial \dot{x}^c}{\partial \dot{x}^c} \right\}
\]

\[
= \frac{g_{ac}\dot{x}^a}{2\sqrt{|g_{ab}\dot{x}^a\dot{x}^b|}} \dot{\theta}_c + \frac{g_{ba}\dot{x}^b}{2\sqrt{|g_{ab}\dot{x}^a\dot{x}^b|}} \dot{\theta}_b = \frac{g_{ac}\dot{x}^a}{2\sqrt{|g_{ab}\dot{x}^a\dot{x}^b|}} + \frac{g_{ba}\dot{x}^b}{2\sqrt{|g_{ab}\dot{x}^a\dot{x}^b|}}
\]

\[
\Rightarrow \frac{\partial F}{\partial \dot{x}^c} = \frac{g_{ac}\dot{x}^a}{2\sqrt{|g_{ab}\dot{x}^a\dot{x}^b|}} + \frac{g_{ba}\dot{x}^b}{2\sqrt{|g_{ab}\dot{x}^a\dot{x}^b|}}
\]

(18.258)

If we replace \( b \) by \( a \) in the first term, we can write

\[
\frac{\partial F}{\partial \dot{x}^c} = \frac{g_{ac}\dot{x}^a}{2\sqrt{|g_{ac}\dot{x}^a\dot{x}^c|}} + \frac{g_{ac}\dot{x}^a}{2\sqrt{|g_{ac}\dot{x}^a\dot{x}^c|}} = \frac{g_{ac}\dot{x}^a}{\sqrt{|g_{ac}\dot{x}^a\dot{x}^c|}} = \frac{g_{ac}\dot{x}^a}{\hat{s}}
\]

so that the Euler-Lagrange Equation becomes

\[
\frac{d}{du} \left( \frac{g_{ac}\dot{x}^a}{\hat{s}} \right) - \frac{\dot{x}^a\dot{x}^b \partial a g_{ab}}{2\hat{s}} = 0.
\]

(18.259)

For the first terms we may write

\[
\frac{d}{du} \left( \frac{g_{ac}\dot{x}^a}{\hat{s}} \right) = \frac{\dot{x}^a \frac{g_{ac}}{\hat{s}}}{du} \frac{du}{\hat{s}} + \frac{g_{ac} \dot{x}^a}{\hat{s}} \frac{d\dot{x}^a}{du} + \frac{g_{ac} \dot{x}^a}{\hat{s}} \frac{d(\frac{1}{\hat{s}})}{du} = \frac{\dot{x}^a \frac{g_{ac}}{\hat{s}}}{du} + \frac{g_{ac} \dot{x}^a}{\hat{s}} \frac{d\dot{x}^a}{du} - \frac{g_{ac} \dot{x}^a}{\hat{s}} \frac{\hat{s}}{\hat{s}^2}
\]

(18.260)

and noting that

\[
\frac{dg_{ac}}{du} = \frac{\partial g_{ac}}{\partial x^b} \frac{dx^b}{du} = (\partial_b g_{ac}) \dot{x}^b
\]

(18.261)

we find

\[
\frac{d}{du} \left( \frac{g_{ac}\dot{x}^a}{\hat{s}} \right) = \frac{1}{\hat{s}} \left[ (\partial_b g_{ac}) \dot{x}^a \dot{x}^b + g_{ac} \dot{x}^a - g_{ac} \frac{\hat{s}}{\hat{s}} \frac{\hat{s}}{\hat{s}} \right]
\]

(18.262)

Now substituting Eq. (18.262) into Eq. (18.259), we find

\[
\frac{1}{\hat{s}} \left[ (\partial_b g_{ac}) \dot{x}^a \dot{x}^b + g_{ac} \dot{x}^a - g_{ac} \frac{\hat{s}}{\hat{s}} \frac{\hat{s}}{\hat{s}} \right] - \frac{\dot{x}^a \dot{x}^b \partial a g_{ab}}{2\hat{s}} = 0.
\]

(18.263)
\[ g_{ac} \dot{x}^a + \frac{1}{2} \left( \partial_b g_{ac} \right) \dot{x}^a \dot{x}^b - \frac{\dot{x}^a \dot{x}^b}{2} \partial_a g_{ab} = \left( \frac{s}{s} \right) g_{ac} \dot{x}^a. \]  

(18.264)

Noting that we can interchange the indices and write

\[ (\partial_b g_{ac}) \dot{x}^a \dot{x}^b = (\partial_a g_{bc}) \dot{x}^b \dot{x}^a \]  

(18.265)

so that

\[ (\partial_b g_{ac}) \dot{x}^a \dot{x}^b = \left( \frac{s}{s} \right) g_{ac} \dot{x}^a. \]  

(18.266)

Substituting Eq. (18.266) into Eq. (18.264), we have

\[ g_{ac} \dot{x}^a + \frac{1}{2} \left[ (\partial_b g_{ac}) \dot{x}^a \dot{x}^b + (\partial_a g_{bc}) \dot{x}^b \dot{x}^a \right] = \left( \frac{s}{s} \right) g_{ac} \dot{x}^a. \]  

(18.267)

and multiplying by \( g^{de} \),

\[ g^{de} g_{ac} \dot{x}^a + \frac{1}{2} g^{de} \left[ (\partial_b g_{ac}) + (\partial_a g_{bc}) - (\partial_a g_{ab}) \right] \dot{x}^a \dot{x}^b = \left( \frac{s}{s} \right) g^{de} g_{ac} \dot{x}^a. \]  

(18.268)

Now applying the relation\( g^{de} g_{ac} = \delta^d_a \), we find

\[ \delta^d_a \dot{x}^a + \frac{1}{2} g^{de} \left[ (\partial_b g_{ac}) + (\partial_a g_{bc}) - (\partial_a g_{ab}) \right] \dot{x}^a \dot{x}^b = \left( \frac{s}{s} \right) \delta^d_a \dot{x}^a. \]  

(18.270)

which simplifies into

\[ \ddot{x}^d + \frac{1}{2} g^{de} \left[ (\partial_b g_{ac}) + (\partial_a g_{bc}) - (\partial_a g_{ab}) \right] \dot{x}^a \dot{x}^b = \left( \frac{s}{s} \right) \dot{x}^d. \]  

(18.271)

Using the expression for the affine connection in terms of the metric

\[ \Gamma^a_{bc} = \frac{1}{2} g^{de} \left[ (\partial_b g_{ac}) + (\partial_a g_{bc}) - (\partial_a g_{ab}) \right] \]  

(18.272)

we find

\[ \ddot{x}^d + \Gamma^a_{bc} \dot{x}^a \dot{x}^b = \left( \frac{s}{s} \right) \dot{x}^d. \]  

(18.273)

Comparing the result in Eq. (18.273) with Eq. (18.252)

\[ \frac{d^2 x^a}{d u^2} + \Gamma^a_{bc} \frac{d x^c}{d u} \frac{d x^b}{d u'} = \left( \frac{d^2 u}{d u' d u} \right) \frac{d x^a}{d u} \]  

(18.274)

we see that these equations are equivalent to one another.
Chapter 19

Tensor Calculus on manifolds

19.1 Tensors fields and rank of a tensor

In order to understand what a tensor is and what is its rank is, it is important to have a better understanding of a vector field, \( \vec{v} \), on a manifold. How do we define a vector field on a manifold. We have learned that a vector field at a given point, \( P \), on a manifold is defined by the tangent plane, \( T_P \), at that point on the manifold. This tangent plane is defined by the tangent vector, \( \vec{t} \). The tangent plane is defined by the basis vectors, \( \hat{e}_a \). We can denote the number produced by the action of the vector, \( \vec{t} \) on the vector, \( \vec{v} \), (the component of the vector field, \( \vec{v} \), on the tangent space) by the scalar product

\[ \vec{t}(\vec{v}) = \vec{t} \cdot \vec{v}. \]

This maps the vector field, \( \vec{v} \), to the tangent space, \( T_p \). In this case one vector \( (\vec{v}) \) is linearly mapped into the tangent space by the tangent vector \( \vec{t} \) (i.e. \( \vec{t} \rightarrow \vec{t}(\vec{v}) \)). Therefore the tangent vector \( \vec{t} \) forms a first rank tensor \( t \).

A tensor: Based on the notion of a vector on a manifold, a tensor is defined by the precise set of operations applied to a set of vectors to produce a scalar and the number of vectors in the set determines the rank of the tensor. If there are \( N \) number of vectors in the set, the tensor is \( N^{th} \) rank tensor (See the table below)

<table>
<thead>
<tr>
<th>Tensor</th>
<th>Operation</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t(\vec{u}, \vec{v}) )</td>
<td>( \vec{u} \cdot \vec{v} )</td>
<td>2</td>
</tr>
<tr>
<td>( t(\vec{u}, \vec{v}, \vec{w}) )</td>
<td>( \vec{u} \cdot \vec{v} \cdot \vec{w} )</td>
<td>3</td>
</tr>
<tr>
<td>( t(\vec{u}, \vec{v}, \vec{w}, \vec{x}) )</td>
<td>( \vec{u} \cdot \vec{v} \cdot \vec{w} \cdot \vec{x} )</td>
<td>4</td>
</tr>
<tr>
<td>( t(\vec{u}, \vec{v}, \vec{w}, \vec{x}, \vec{y}) )</td>
<td>( \vec{u} \cdot \vec{v} \cdot \vec{w} \cdot \vec{x} \cdot \vec{y} )</td>
<td>5</td>
</tr>
</tbody>
</table>

From this table we can easily see that a scalar field can be classified as a zero ranked tensor field since it does not depend on a vector field.
A tensor is a linear map of the vectors into the real and therefore any ranked tensor is linear. This means, for example, for 1st rank tensor, we must have
\[ \bar{t}(\alpha \vec{u} + \beta \vec{v}) = \alpha \bar{t}(\vec{u}) + \beta \bar{t}(\vec{v}) \] (19.2)

**Homework:** Show that 2nd rank tensor is linear.

**Components of a tensor:** We recall that the tangent plane is defined in terms of the basis ($\vec{e}_a$) or dual basis vectors ($\vec{e}^b$). When vectors are expressed in terms of basis or dual basis vectors we can determine the components of a tensor in different forms. But first let's consider if the vectors, $\vec{v}$ and $\vec{u}$, are just the basis or the dual basis vectors. In this case we have

(a) 1st rank tensor
\[
\begin{align*}
t(\vec{e}_a) &= \bar{t} \cdot \vec{e}_a = t_a, \\
t(\vec{e}^b) &= \bar{t} \cdot \vec{e}^b = t^b.
\end{align*}
\] (19.3) \hspace{1cm} (19.4)

(b) 2nd rank tensor
\[
\begin{align*}
t(\vec{e}_a, \vec{e}_b) &= t^{ab}, \\
t(\vec{e}^a, \vec{e}^b) &= t^a_b.
\end{align*}
\] (19.5) \hspace{1cm} (19.6)

**Example 4.1** Let’s reconsider the 2D sphere in the 3D manifold. For the vector on the tangent plane shown in Fig. (fig4.000)

\[ t(\vec{e}_a) = \vec{e}_1 + \vec{e}_2 \text{ or } t(\vec{e}_a) = \vec{e}^1 + \vec{e}^2 \]
we have
\[ t_1 = t(\vec{e}_a) \cdot \vec{e}_1, \quad t_2 = t(\vec{e}_a) \cdot \vec{e}_2 \Rightarrow t(\vec{e}_a) = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \]
\[ \Rightarrow t(\vec{e}_a) = \begin{pmatrix} \vec{e}_1 \cdot \vec{e}_1 + \vec{e}_2 \cdot \vec{e}_1 \\ \vec{e}_1 \cdot \vec{e}_2 + \vec{e}_2 \cdot \vec{e}_2 \end{pmatrix} = \begin{pmatrix} g_{11} + g_{21} \\ g_{12} + g_{22} \end{pmatrix} \]
and we recall for 2D sphere \( g_{21} = g_{12} = 0 \),
\[ t(\vec{e}_a) = \begin{pmatrix} g_{11} \\ g_{22} \end{pmatrix} \]
which is a 1-st rank tensor. Now let’s consider a quantity defined by by
an operation on set that consist of two vectors as
\[ t(\vec{e}_a, \vec{e}_b) = (\vec{e}_1 \cdot \vec{e}_1) + (\vec{e}_2 \cdot \vec{e}_1) + (\vec{e}_1 \cdot \vec{e}_2) + (\vec{e}_2 \cdot \vec{e}_2) \]
\[ \Rightarrow t(\vec{e}_a, \vec{e}_b) = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \]
which forms a 2-nd rank tensor. Using the dual basis vector, we can also
express this 2-nd rank tensor as
\[ t(\vec{e}_a, \vec{e}^b) = (\vec{e}_1 \cdot \vec{e}^d) + (\vec{e}_2 \cdot \vec{e}^d) + (\vec{e}_1 \cdot \vec{e}^2) + (\vec{e}_2 \cdot \vec{e}^2) \]
\[ \Rightarrow t(\vec{e}_a, \vec{e}^b) = \begin{pmatrix} g^1_1 & g^1_2 \\ g^2_1 & g^2_2 \end{pmatrix} \]
or
\[ t(\vec{e}^a, \vec{e}^b) = (\vec{e}^a \cdot \vec{e}^d) + (\vec{e}^a \cdot \vec{e}^d) + (\vec{e}^2 \cdot \vec{e}^2) + (\vec{e}^2 \cdot \vec{e}^2) \]
\[ \Rightarrow t(\vec{e}^a, \vec{e}^b) = \begin{pmatrix} g^1_1 & g^1_2 \\ g^2_1 & g^2_2 \end{pmatrix} \]
Now we can apply the linearity of tensors to determine the components of
a tensor for the general case where we have two vectors, \( \vec{v} \) and \( \vec{u} \), expressed in
terms of the covariant or contravariant components in terms of the basis
and dual basis vectors as
\[ \vec{u} = u_a \vec{e}^a, \quad \vec{v} = v_b \vec{e}^b, \quad (19.7) \]
\[ \vec{\bar{u}} = u^a \vec{e}_a, \quad \vec{\bar{v}} = v^b \vec{e}_b. \quad (19.8) \]
(a) 1\textsuperscript{st} rank tensor
\[ t(\vec{u}) = t(u^a \vec{e}^a) = t(\vec{e}^a) u_a = t^a u_a, \quad (19.9) \]
\[ t(\vec{\bar{u}}) = t(u^a \vec{e}_a) = u^a t(\vec{e}^a) = t_a u^a \quad (19.10) \]
(b) 2\textsuperscript{nd} rank tensor
\[ t(\vec{u}, \vec{v}) = t(u^a \vec{e}_a, v_b \vec{e}^b) = t(\vec{e}_a, \vec{e}^b) u^a v_b = t_a^b u^a v_b, \quad (19.11) \]
\[ t(\vec{\bar{u}}, \vec{\bar{v}}) = t(u^a \vec{e}_a, v^b \vec{e}_b) = t(\vec{e}^a, \vec{e}^b) u_a v^b = t^a_b u_a v^b, \quad (19.12) \]
\[ t(\vec{\bar{u}}, \vec{v}) = t(u^a \vec{e}_a, v^b \vec{e}^b) = t(\vec{e}^a, \vec{e}_b) u^a v^b = t_{ab} u^a v^b. \quad (19.13) \]
Symmetries of a tensor: A tensor can be symmetric or antisymmetric. For a second ranked tensor $t(\vec{u}, \vec{v})$

$$t(\vec{u}, \vec{v}) = \begin{cases} -t(\vec{v}, \vec{u}) & \text{Antisymmetric} \\ t(\vec{v}, \vec{u}) & \text{symmetric} \end{cases}$$

Any tensor can be expressed as a sum of symmetric and antisymmetric tensor. Again if we consider a second rank tensor with elements $t_{ab}$, we can express this elements as

$$t_{ab} = \frac{1}{2} (t_{ab} + t_{ba}) + \frac{1}{2} (t_{ab} - t_{ba}). \quad (19.14)$$

Introducing the notations for the symmetric part

$$t_{(ab)} = \frac{1}{2} (t_{ab} + t_{ba}) \quad (19.15)$$

and the antisymmetric part

$$t_{[ab]} = \frac{1}{2} (t_{ab} - t_{ba}) \quad (19.16)$$

we can write

$$t_{ab} = t_{(ab)} + t_{[ab]} \quad (19.17)$$

$N^{th}$ ranked tensor: for an $N^{th}$ ranked tensor the symmetric and antisymmetric covariant components of the tensor, $t_{a_1a_2...a_N}$ are given by

$$t_{(a_1a_2...a_N)} = \frac{1}{N!} (\text{addition over all permutations of the indices } a_1a_2...a_N) \quad (19.18)$$

and

$$t_{[a_1a_2...a_N]} = \frac{1}{N!} (\text{Alternating subtraction and addition over all permutations of the indices } a_1a_2...a_N) \quad (19.19)$$

$$t_{[a_1a_2a_3]} = \frac{1}{3!} (t_{a_1a_2a_3}) \quad (19.20)$$

For example for 3rd rank tensor, we have

$$t_{(a_1a_2a_3)} = \frac{1}{3!} (t_{a_1a_2a_3} + t_{a_2a_1a_3} + t_{a_1a_2a_3} + t_{a_3a_2a_1} + t_{a_1a_2a_3} + t_{a_3a_2a_3}) \quad (19.21)$$

and

$$t_{(a_1a_2a_3)} = \frac{1}{3!} (t_{a_1a_2a_3} - t_{a_2a_1a_3} + t_{a_1a_2a_3} - t_{a_3a_2a_1} + t_{a_1a_2a_3} - t_{a_3a_2a_3}) \quad (19.22)$$

Particular subset of indices permutation: we have a different notations when the permutation is to particular subset of indices. This is described using a 4th
19.1. TENSORS FIELDS AND RANK OF A TENSOR

rank tensor.

\[
t_{(ab)cd} = \frac{1}{2} (t_{abcd} + t_{bacd})
\]

(19.23)
symmetric permutation to indices a and b only

\[
t_{(ab)cd} = \frac{1}{2} (t_{abcd} - t_{bacd})
\]

(19.24)
antisymmetric permutation to indices a and b only

\[
t_{a[b|c|d]} = \frac{1}{2} (t_{abcd} - t_{bdaa})
\]

(19.25)
antisymmetric permutation to indices b and d only

Note that the symbol \( || \) are used to exclude unwanted indices from the symmetrization ( ) antisymmetrization \( [ ] \) implied.

**Example 4.2** Let’s reconsider the 3D sphere in the 4D manifold. We recall the tangent space at a point on this 3D sphere form three basis and dual basis vectors, \( \mathbf{e}_1, \mathbf{e}_2, \) and \( \mathbf{e}_3 \) (or \( \mathbf{\bar{e}}_r, \mathbf{\bar{e}}_\theta, \) and \( \mathbf{\bar{e}}_\phi \)). Write the symmetric, \( t_{(123)} \) and antisymmetric components \( t_{(123)} \)

**Homework:** Find the explicit form of \( t_{(ab)(cd)} \)

The metric tensor revisited: A good example of a 2\(^{nd}\) rank tensor is the metric tensor, \( g \), that we already have seen. Here more generally we define the metric tensor as a linear map of two vectors \( \vec{u} \) and \( \vec{v} \) into the number that is the inner product

\[
g(\vec{u}, \vec{v}) = \vec{u} \cdot \vec{v}
\]

(19.26)
We can easily see that the metric tensor is symmetric tensor as
\[ g(\vec{u}, \vec{v}) = g(\vec{v}, \vec{u}) = \vec{v} \cdot \vec{u} \quad (19.27) \]

We recall that covariant and contravariant components of the metric tensor are given by
\[ g_{ab} = g(\vec{e}_a, \vec{e}_b) = \vec{e}_a \cdot \vec{e}_b, \quad g^{ab} = g(\vec{e}^a, \vec{e}^b) = \vec{e}^a \cdot \vec{e}^b \quad (19.28) \]
and the mixed components
\[ g^{b}_a = g(\vec{e}_a, \vec{e}^b) = g(\vec{e}^a, \vec{e}_b) = \delta^a_b \quad (19.29) \]

Raising and lowering tensor indices. Consider the third rank tensor, \( t \), expressed in terms of its covariant components
\[ t_{abc} = \vec{e}_a \cdot \vec{e}_b \cdot \vec{e}_c \quad (19.30) \]
or mixed components
\[ t'_{ab} = \vec{e}_a \cdot \vec{e}_b \cdot \vec{e}^c \quad (19.31) \]
We recall from chapter 3 that the covariant form of the metric can be used to lower and the contravariant can be used to raise the index. Thus to lower the index in Eq. (19.31) we multiply by the covariant form of the metric tensor, \( g_{dc} \),
\[ g_{dc} t'_{ab} = g_{dc} \vec{e}_a \cdot \vec{e}_b \cdot \vec{e}^c \quad (19.32) \]
noting that
\[ g_{dc} = \vec{e}_d \cdot \vec{e}_c, \quad \text{and} \quad \vec{e}_b \cdot \vec{e}^c = \delta^b_c \]
we have
\[ g_{dc} t'_{ab} = \vec{e}_d \cdot \vec{e}_c \cdot \vec{e}_a \cdot \delta^c_b = \vec{e}_d \cdot \vec{e}_b \cdot \vec{e}_a = t_{dab} \quad (19.33) \]
so that
\[ t_{dab} = g_{dc} t'_{ab} \quad (19.34) \]
Since the indices are dummy variables and tensor \( t_{abc} \) is symmetric, one can write
\[ t_{abc} = g_{dc} t'_{dc} \quad (19.35) \]

**Homework:** Raise the tensor \( t_{abc} \) to \( t'_{bc} \).

### 19.2 Mapping tensors into tensors

We have learned that tensor maps vectors into real numbers. We have seen that 1\textsuperscript{st} rank tensor \( t(\vec{u}) \) maps the vector \( \vec{u} \) into real numbers in the tangent space.
\[ t(\vec{u}) = t(u^a \hat{e}_a) = u^a t(\hat{e}_a) = t_a u^a \quad (19.36) \]
and the second rank tensor \( t(\vec{u}, \vec{v}) \) maps the vectors \( \vec{u} \) and \( \vec{v} \) into real numbers in the tangent space
\[ t(\vec{u}, \vec{v}) = t(u^a \hat{e}_a, v^b \hat{e}_b) = t(\hat{e}_a, \hat{e}_b) u^av^b = t_{ab} u^av^b. \]
Now the question is can we map a tensor into another tensor of a different rank. Consider a third rank tensor $t(\vec{u}, \vec{v}, \vec{w})$ which maps the three vectors into real numbers. Now instead if we replace the two vectors $(\vec{u}, \vec{v})$ by the basis vectors $(\vec{e}_a, \vec{e}_b)$ on the tangent plane, we have

$$t(\vec{e}_a, \vec{e}_b, \vec{w}) = \vec{e}_a \cdot \vec{e}_b \cdot \vec{w}, \quad (19.37)$$

so that if we express the vector, $\vec{w}$, in terms of its controvariant components $\vec{w} = w^c \vec{e}_c$ (19.38)

we find

$$t(\vec{e}_a, \vec{e}_b, \vec{w}) = \vec{e}_a \cdot \vec{e}_b \cdot w^c \vec{e}_c = (\vec{e}_a \cdot \vec{e}_b \cdot \vec{e}_c) w^c = t_{abc}w^c = s_{ab}. \quad (19.39)$$

which is a 2-nd rank tensor. We see that the 3-rd rank tensor $t$ maps the vector, $\vec{w}$, into a second rank tensor $s$.

As another example let’s consider the third rank tensor $t(\vec{v}, \vec{e}_b, \vec{w}) = \vec{v} \cdot \vec{e}_b \cdot \vec{w}$ (19.40)

Suppose we express the vectors, $\vec{v}$ and $\vec{w}$, in terms of its controvariant components $\vec{v} = v^a \vec{e}_a$, $\vec{w} = w^c \vec{e}_c$, then

$$t(\vec{e}_a, \vec{e}_b, \vec{w}) = v^a \vec{e}_a \cdot \vec{e}_b \cdot w^c \vec{e}_c = (\vec{e}_a \cdot \vec{e}_b \cdot \vec{e}_c) v^a w^c = t_{abc}v^a w^c = v_b, \quad (19.41)$$

where we find a 1-st rank tensor, $s_{b}$. This means the 3rd rank tensor mapped the two vectors into 1st rank tensor (a vector).

### 19.3 Elementary tensor operations

**Adding, subtracting, and multiplying by a scalar:**

$$s_{ab} = \alpha s(\vec{e}_a, \vec{e}_b) = \alpha t(\vec{e}_a, \vec{e}_b) = \alpha t_{ab} = \alpha r_{ab}. \quad (19.43)$$

**Outer product:** Consider two first rank tensors (two vectors) $u(\vec{p})$ and $v(\vec{q})$. The outer product or (tensor product) of these two tensors, which is denoted by $u \otimes v$ is given by

$$u \otimes v(\vec{p}, \vec{q}) = u(\vec{p})v(\vec{q}), \quad (19.44)$$

which is a second rank tensor. Suppose $(\vec{p}, \vec{q}) \rightarrow (\vec{e}_a, \vec{e}_b)$, then

$$u \otimes v(\vec{e}_a, \vec{e}_b) = u(\vec{e}_a)v(\vec{e}_b) = u_av_b. \quad (19.45)$$

Now let’s consider 2-nd rank tensor, $t(\vec{p}, \vec{q})$, and 1-st rank tensor (a vector), $s(\vec{r})$, the outer product of these two tensors give another tensor, $h$, of a different rank given by

$$t \otimes s(\vec{p}, \vec{q}, \vec{r}) = t(\vec{p}, \vec{q})s(\vec{r}) = h. \quad (19.46)$$
The tensor $h$ is a 3-rd rank tensor. Using the basis and dual basis vectors we may express the components as

\[
t \otimes s (e_a, e_b, e_c) = t (e_a, e_b) s (e_c) = h_{abc} \quad (19.47)
\]

\[
t \otimes s (e^a, e_b, e_c) = t (e^a, e_b) s (e_c) = h^b_{ac} \quad (19.48)
\]

\[
t \otimes s (e^a, e^b, e_c) = t (e^a, e^b) s (e_c) = h^c_{ab} \quad (19.49)
\]

\[
t \otimes s (e_a, e_b, e^c) = t (e_a, e_b) s (e^c) = h^c_{ab} \quad (19.50)
\]

**Example 4.2** Let's consider two vectors in the tangent space for a point in a 3D sphere embedded in 4D manifold given by

\[
\vec{p} = p^a e_a, \quad \vec{q} = q^b e_b
\]

where $a, b = 1, 2, 3$, and $\vec{e}_1, \vec{e}_2$, and $\vec{e}_3$ (or $\vec{e}_r, \vec{e}_\theta$, and $\vec{e}_\phi$). Find the components of the 2-nd rank tensor for

\[
u \otimes v (\vec{p}, \vec{q}) = u (\vec{p}) v (\vec{q}), \quad (19.51)
\]

**Sol:** The components of this 2-nd rank tensor are given by

\[
u \otimes v (\vec{p}, \vec{q}) = p^a q^b e_a e_b
\]

\[
= p^1 q^1 e_1 e_1 + p^1 q^2 e_1 e_2 + p^1 q^3 e_1 e_3
+ p^2 q^1 e_2 e_1 + p^2 q^2 e_2 e_2 + p^2 q^3 e_2 e_3
+ p^3 q^1 e_3 e_1 + p^3 q^2 e_3 e_2 + p^3 q^3 e_3 e_3
\]

and using a matrix this is expressed as

\[
u \otimes v (\vec{p}, \vec{q}) = \begin{bmatrix}
p^1 q^1 & p^1 q^2 & p^1 q^3 \\
p^2 q^1 & p^2 q^2 & p^2 q^3 \\
p^3 q^1 & p^3 q^2 & p^3 q^3
\end{bmatrix}
\]

**NB:** From what we learned in all mathematical or physics courses up to this point, what we know is that inner product of two 1-st rank tensor (two vectors) is commutative. However, that generalization does not apply to higher ranked tensor. So from now on we must keep in mind that inner product of tensors in general is not commutative including the 1-st rank tensor for the reason described in terms of tensor contraction. For example, if $t$ is a 2-nd rank tensor and $s$ is a first rank tensor, the inner product

\[
t \cdot s = t^{ab} s_b \quad (19.52)
\]

and

\[
s \cdot t = t^{ab} s_a \quad (19.53)
\]

are not necessarily the same.
Like vectors (1-st rank tensor) tensors of higher rank are geometrical objects too: We already know that vectors are geometrical objects that can be made up from a linear combination of the basis vectors

\[ t = t_a e^a = t^a e_a. \] (19.54)

The vector that defines a given geometry on a manifold, does not depend on how it is represented. The geometry that a vector defines remain the same geometry whatever representation we used to describe the vector. The same is true for higher ranked tensors. As an example, let’s consider a 2-nd rank tensor with component \( t_{ab} \) constructed from the outer product of two basis vectors

\[ t = e_a \otimes e_b. \] (19.55)

The controvariant components of \( t \) can then be expressed as

\[ t = e_a \otimes e_b (e^c, e^d) = \delta_a^c \delta_b^d. \] (19.56)

Now suppose we have some general 2-nd rank tensor, \( t \), whose controvariant components are \( t^{ab} (e_a \otimes e_b) \). The action of this second rank tensor on two basis vectors

\[ t^{ab} (e_a \otimes e_b) (e^c, e^d) = t^{cd} \] (19.57)

Therefore what is true for example for 2-d rank tensor

\[ t = t^{ab} (e_a \otimes e_b) = t^b_a (e^a \otimes e_b) = t^a_b (e_a \otimes e^b) \] (19.58)

is true for any ranked tensor.

### 19.4 Tensors and coordinate transformations

We recall the coordinates basis and the dual basis vectors, under the coordinate transformation \( x^a \to x'^a \), are transformed as

\[ e'_c = \frac{\partial x^a}{\partial x'^c} e_a. \] (19.59)

and

\[ e'^a = \frac{\partial x'^a}{\partial x^c} e^c. \] (19.60)

**1-st rank tensor**: Suppose we have a 1-st rank tensor, \( t \)

\[ t (e^a) = t_a e^a = t^a e_a. \] (19.61)

in the \( x'^a \) coordinate system is given by

\[ t'^a = t' (e'^a) = \frac{\partial x'^a}{\partial x^c} t (e^c) = \frac{\partial x'^a}{\partial x^c} t^c \] (19.62)

\[ t'_a = t' (e'_a) = \frac{\partial x^c}{\partial x'^a} t (e_c) = \frac{\partial x^c}{\partial x'^a} t_c \] (19.63)
2-nd rank tensor: Suppose we have a 2-nd rank tensor, \( t(e^a, e^b) \), and \( t(e_a, e_b) \) in the \( x^a \) coordinate system is given by
\[
 t' (e^a, e^b) = \frac{\partial x^a}{\partial x^c} \frac{\partial x^b}{\partial x^d} t^{cd} \tag{19.64}
\]
\[
 t' (e'_a, e'_b) = \frac{\partial x^c}{\partial x^a} \frac{\partial x^d}{\partial x^b} t_{cd} \tag{19.65}
\]
\[
 t' (e'_a, e'_b) = \frac{\partial x^c}{\partial x^a} \frac{\partial x^d}{\partial x^b} t_{cd} \tag{19.66}
\]

3-nd rank tensor: Suppose we have a mixed 3-rd rank tensor, \( t^{ab} \) \( \rightarrow \) \( t(e_a, e_b, e^c) \) in the \( x^a \) coordinate system is given by
\[
 t (e'_a, e'_b, e'^c) = \frac{\partial x^c}{\partial x^a} \frac{\partial x^d}{\partial x^b} \frac{\partial x^e}{\partial x^f} t^{cd} \tag{19.67}
\]

19.5 Tensor equations and the quotient theorem

A tensor equation which holds in one coordinate system must hold in another coordinate system. Suppose we have an equation that states two second rank tensors, \( t \) and \( s \) are equal in the \( x^a \) coordinate system. That means
\[
 t_{cd} = s_{cd} \tag{19.68}
\]

Multiplying both sides of this equation by
\[
 \frac{\partial x^e}{\partial x^a} \frac{\partial x^d}{\partial x^b}
\]
we have
\[
 \frac{\partial x^e}{\partial x^a} \frac{\partial x^d}{\partial x^b} t_{cd} = \frac{\partial x^e}{\partial x^a} \frac{\partial x^d}{\partial x^b} s_{cd} \tag{19.69}
\]
so that using Eq. (19.66), we find
\[
 t'_{ab} = s'_{ab} \tag{19.70}
\]

Well Eq. (19.70) shows that the equality holds under the coordinate transformation. However, the question is that these components (set of quantities) are actually form a tensor. The quotient theorem set’s the condition for a set of quantities actually represent a tensor component.

The quotient theorem: if a set of quantities when contracted with a tensor produces another tensor, then the original set of quantities are also a tensor.

Suppose in an \( N \) dimensional manifold you are given a 3-rd rank tensor, \( t \), and 1-st rank tensor, \( v \). The tensor, \( t \), has a set of \( N^3 \) quantities \( t_{ab} \), and the tensor \( v \) has \( N \) quantities of \( v^a \). Suppose we form a set of \( N^2 \) quantities by contracting the 4-th rank tensor formed by the outer product of these two
tensors (i.e. \( s^a_b = t^c_b v^c \)). Under coordinate transformation \( x^a \rightarrow x'^a \), these set of elements, using the transformation relations,

\[
t^a = \frac{\partial x'^a}{\partial x^c} v^c
\]
(19.71)

\[
t'_a = \frac{\partial x^c}{\partial x'^a} t^c
\]
(19.72)
in the new coordinate system are given by

\[
s'^a_b = \frac{\partial x'^a}{\partial x^d} s^d \frac{\partial x^e}{\partial x'^b} s^e = \frac{\partial x^a}{\partial x'^d} \frac{\partial x^e}{\partial x'^b} s^d s^e = t^a_b v^c = \frac{\partial x^a}{\partial x'^d} \frac{\partial x^e}{\partial x'^b} t^d e f v^f.
\]
(19.73)

In the relation

\[
v'^f = \frac{\partial x'^f}{\partial x'^e} v^e
\]
(19.74)

by switching the the primes with the none primes, we can write

\[
v^f = \frac{\partial x^f}{\partial x^e} v^e
\]
(19.75)

Substituting this expression into Eq. (19.73), we find

\[
t^a_b v'^c = \frac{\partial x'^a}{\partial x^d} \frac{\partial x^e}{\partial x'^b} t^d e f v^f \Rightarrow \left( t^a_b v^c - \frac{\partial x'^a}{\partial x^d} \frac{\partial x^e}{\partial x'^b} \frac{\partial x^f}{\partial x'^e} t^d e f \right) v^c = 0
\]
(19.76)

There follows that for an arbitrary vector components, \( v^c \),

\[
t^a_b = \frac{\partial x'^a}{\partial x^e} \frac{\partial x^f}{\partial x'^b} \frac{\partial x^d}{\partial x^c} t^d e f.
\]
(19.77)

We made a contraction to the set of elements \( t^a_b \) and still we found a tensor. Therefore the set of element \( t^a_b \) must be the component of the 3-rd rank tensor. That is the quotient theorem.

### 19.6 Covariant derivatives of a tensor

Suppose we consider the 1-st rank tensor \( v^a \) in the \( x^a \) coordinates. The derivative of this tensor can be expressed as

\[
\frac{\partial v^a}{\partial x'^b} = \frac{\partial v^a}{\partial x^c} \frac{\partial x^c}{\partial x'^b} = \frac{\partial x^c}{\partial x'^b} \frac{\partial v^a}{\partial x^c}
\]
(19.78)

Using the transformation for 1-st rank tensor, we have

\[
v'^a = \frac{\partial x'^a}{\partial x^d} v^d
\]
(19.79)

so that

\[
\frac{\partial v'^a}{\partial x'^b} = \frac{\partial x^c}{\partial x'^b} \frac{\partial x'^a}{\partial x^c} = \frac{\partial x^c}{\partial x'^b} \frac{\partial x'^a}{\partial x^d} \frac{\partial x^d}{\partial x'^c} + \frac{\partial x^c}{\partial x'^b} \frac{\partial x'^a}{\partial x^d} \frac{\partial x^d}{\partial x'^c}
\]
(19.80)
We recall the covariant derivative from chapter 3
\[ \nabla_b v^a = \partial_b v^a + \Gamma^a_{cb} v^c \]  
(19.81)
so that the gradient of the first rank tensor can be expressed as
\[ \nabla v = (\nabla_b v^a) e_a = \partial_b v^a + \Gamma^a_{cb} v^c \]  
(19.82)
\( \nabla_b v^a \) forms a mixed 2-nd rank tensor. We denote this second rank tensor by \( \nabla^a v \). Noting that
\[ \nabla v = e^a \partial_a, v = v^b e_b \]
we can express
\[ \nabla v = e^a \partial_a \otimes v^b e_b = e^a \otimes \partial_a (v^b e_b) = (\nabla_a v^b) e^a \otimes e_b \]  
(19.83)
Let’s consider the the covariant derivative of a second rank tensor \( t \) with expressed in terms of its controvariant components \( t^{ab} \)
\[ \nabla_c t = \nabla_c t^{ab} e_a \otimes e_b \]  
(19.84)
Using the product rule, we have
\[ \partial_c t = (\partial_c t^{ab}) e_a \otimes e_b + t^{ab} (\partial_c e_a) \otimes e_b + t^{ab} e_a \otimes (\partial_c e_b) \]  
(19.85)
and recalling that
\[ \partial_c \hat{e}_b = \Gamma^f_{bc} \hat{e}_f \]  
(19.86)
one can write
\[ \partial_c t = (\partial_c t^{ab}) e_a \otimes e_b + t^{ab} \Gamma^f_{ca} \hat{e}_f \otimes e_b + t^{ab} e_a \otimes (\partial_c e_b) \]  
(19.87)
\[ \partial_c t = (\partial_c t^{ab}) (e_a \otimes e_b) + t^{ab} \Gamma^f_{ca} (\hat{e}_f \otimes e_b) + t^{ab} \Gamma^f_{cb} (e_a \otimes \hat{e}_f) \]  
(19.88)
If we interchange the indices \( f \) and \( a \) in the second term and \( f \) by \( b \) in the third terms, we have
\[ \partial_c t = (\partial_c t^{ab}) (e_a \otimes e_b) + t^{fb} \Gamma^a_{cf} (\hat{e}_a \otimes e_b) + t^{ab} \Gamma^b_{cf} (e_a \otimes \hat{e}_b) \]  
(19.89)
which can be rewritten as
\[ \partial_c t = [(\partial_c t^{ab} + \Gamma^a_{cf} t^{fb} + \Gamma^b_{cf} t^{af})] (e_a \otimes e_b) = (\nabla_c t^{ab}) \hat{e}_a \otimes e_b \]  
(19.90)
where
\[ \nabla_c t^{ab} = \partial_c t^{ab} + \Gamma^a_{cd} t^{db} + \Gamma^b_{cd} t^{ad} \]  
(19.91)
is the covariant derivative.

**Homework:** Show that for the covariant derivatives of the mixed and covariant component of a 2-nd rank tensor \( t \) are given by
\[ \nabla_c t^{ab} = \partial_c t^{ab} + \Gamma^a_{db} t^{c} - \Gamma^b_{dc} t^{da} \]  
(19.92)
\[ \nabla_c t^{ab} = \partial_c t^{ab} - \Gamma^d_{ac} t^{db} - \Gamma^d_{bc} t^{ad} \]  
(19.93)
Useful relation

\[ \partial_c e^a = -\Gamma^a_{bc} e^b \]  

(19.94)

**Homework:** Show that the covariant derivative of the metric tensor is zero

\[ \nabla g = 0 \]

Suppose we represent the metric tensor in terms of its contravariant components, \( g^{ab} \), then you must show that the covariant derivative expressed as

\[ \nabla_c g^{ab} = \partial_c g^{ab} + \Gamma^a_{cd} g^{db} + \Gamma^b_{cd} g^{ad} = 0 \]  

(19.95)

Useful relations, for example, the affine connection and the metric are related by

\[ \Gamma^f_{bc} = \frac{g^{fd}}{2} (\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc}) \]  

(19.96)

Application of the property of the metric tensor: Suppose we have a 2nd rank tensor, \( t \), for which we want to find the covariant derivative from its components, for example, in mixed form. That means we want to find \( \nabla_c t^{ab} \) from \( t_a^b \). We can express the covariant components of this tensor using contraction as

\[ t_{ab} = g^{bd} t_a^d. \]

Note that the metric tensor is symmetric (\( g^{bd} = g^{db} \)). Then

\[ t_{ab} = g^{bd} t_a^d \]

\[ \nabla_c t^{ab} = \nabla_c (g^{bd} t_a^d) = (\nabla_c g^{bd}) t_a^d + g^{bd} \nabla_c t_a^d \]

(19.97)

since

\[ \nabla_c g^{bd} = 0 \]

we can express

\[ \nabla_c t^{ab} = g^{bd} \nabla_c t_a^d. \]

(19.98)

### 19.7 Intrinsic derivative

Like vectors (1-rank tensor), tensor of rank 2 or more can depend on a submanifold instead of the entire manifold. For example, a given tensor \( t \) can depend on a curve \( C \) on the manifold that defined by some parameter \( u \). This curve may be defined by the equation \( x^a(u) \). The tensor, \( t \), expressed in terms of its contravariant components, then in terms of this parameter that the curve, \( C \), depends on can be expressed as

\[ t(u) = t^{ab} e_a(u) \otimes e_b(u). \]

The intrinsic derivative will then be

\[ \frac{dt(u)}{du} = \frac{d}{du} \left[ t^{ab} e_a(u) \otimes e_b(u) \right] \]

\[ = \frac{dt^{ab}}{du} e_a(u) \otimes e_b(u) + t^{ab} \frac{de_a(u)}{du} \otimes e_b(u) + t^{ab} e_a(u) \otimes \frac{de_b(u)}{du} \]
In terms of the coordinates \( x^a (u) \), we can write
\[
\frac{de_a (u)}{du} = \frac{dx^c}{du} \frac{de_b (u)}{du} = \frac{dx^b}{du}
\]
and using the affine connection
\[
\partial_c e_b = \frac{de_a (u)}{dx^c} = \Gamma^f_{bc} \hat{e}_f
\]
we have
\[
\frac{de_a (u)}{du} = \Gamma^f_{a c} \hat{e}_f (u) \frac{dx^c}{du}, \quad \frac{de_b (u)}{du} = \Gamma^f_{b c} \hat{e}_f (u) \frac{dx^b}{du}
\]
so that
\[
\frac{dt (u)}{du} = \frac{dt^{ab}}{du} e_a (u) \otimes e_b (u) + t^{ab} \Gamma^f_{ac} \hat{e}_f (u) \frac{dx^c}{du} \otimes e_b (u) + t^{ab} \Gamma^f_{bc} \hat{e}_f \frac{dx^b}{du}
\]
(19.100)

Replacing the dummy index \( a \) by \( f \) in the first term, we can write
\[
\frac{dt (u)}{du} = \frac{dt^{f b}}{du} e_f (u) \otimes e_b (u) + t^{ab} \Gamma^f_{ac} \hat{e}_f (u) \frac{dx^c}{du} \otimes e_b (u) + t^{ab} \Gamma^f_{bc} \hat{e}_f \frac{dx^b}{du}
\]
(19.101)

Noting that by making the following dummy index change \( b \rightarrow d \) followed by \( a \rightarrow b \)
\[
t^{ab} \Gamma^f_{bc} \frac{dx^b}{du} e_a (u) \otimes e_f (u) = t^{ad} \Gamma^f_{ad} \frac{dx^d}{du} e_a (u) \otimes e_f (u) = t^{bd} \Gamma^f_{bd} \frac{dx^b}{du} e_a (u) \otimes e_f (u)
\]
(19.102)

we find
\[
\frac{dt (u)}{du} = \left[ \frac{dt^{f b}}{du} + t^{ab} \Gamma^f_{ac} \frac{dx^c}{du} + t^{bd} \Gamma^f_{bd} \frac{dx^d}{du} \right] e_f (u) \otimes e_b (u)
\]
(19.104)

that we expressed as
\[
\frac{dt (u)}{du} = \frac{Dt^{f b}}{Du} e_f (u) \otimes e_b (u)
\]
(19.105)

where
\[
\frac{Dt^{f b}}{Du} = \frac{dt^{f b}}{du} + t^{ab} \Gamma^f_{ac} \frac{dx^c}{du} + t^{bd} \Gamma^f_{bd} \frac{dx^d}{du}
\]
(19.106)

is called the intrinsic (absolute) derivative of the component \( t^{f b} \) along the curve defined by \( x^a (u) \). For the sake of convenience we make change of dummy indices \( (f \rightarrow a) \) in Eq. (19.105) and \( (a \rightarrow d) \) in Eq. (19.106) as
\[
\frac{dt (u)}{du} = \frac{Dt^{ab}}{Du} e_a (u) \otimes e_b (u)
\]
(19.107)
where
\[
\frac{Dt^{ab}}{Du} = \frac{dt^{ab}}{du} + t^{db} \Gamma^a_{dc} \frac{dx^c}{du} + t^{bc} \Gamma^a_{cd} \frac{dx^c}{du}.
\] (19.108)

We can switch the dummy indices \(c\) and \(d\) in the third term as the affine is symmetric for Torsionless
\[
\frac{Dt^{ab}}{Du} = \frac{dt^{ab}}{du} + t^{db} \Gamma^a_{dc} \frac{dx^c}{du} + t^{bc} \Gamma^a_{cd} \frac{dx^c}{du}.
\]

It can be easily shown that
\[
\frac{dt}{Du} = \frac{Dt^{ab}}{Du} e^a (u) \otimes e^b (u) = \frac{Dt^{ab}}{Du} e^a (u) \otimes e^b (u) = \frac{Dt^{ab}}{Du} e^a (u) \otimes e^b (u).
\] (19.109)

Like vectors parallel-transported, we can say tensors are parallel-transported when
\[
\frac{Dt^{ab}}{Du} = 0.
\] (19.110)

Suppose we pretend that the tensor depends on the entire manifold instead of submanifold defined by some curve \(C\), we can write
\[
\frac{dt^{ab}}{du} = \frac{\partial t^{ab}}{\partial x^c} \frac{dx^c}{du} = \partial_c t^{ab} \frac{dx^c}{du}
\]
so that the intrinsic derivative can be expressed as
\[
\frac{Dt^{ab}}{Du} = \partial_c t^{ab} \frac{dx^c}{du} + t^{db} \Gamma^a_{dc} \frac{dx^c}{du} + t^{bc} \Gamma^a_{cd} \frac{dx^c}{du} = \left( \partial_c t^{ab} + t^{db} \Gamma^a_{dc} + t^{bc} \Gamma^a_{cd} \right) \frac{dx^c}{du}
\]

Using the result in Eq. (19.91) we can write
\[
\frac{Dt^{ab}}{Du} = \nabla_c t^{ab} \frac{dx^c}{du}
\] (19.111)