মানুষের জ্ঞান ও ভাবকে বইয়ের মধ্যে সঞ্চিত করিবার যে একটা প্রচুর সুবিধা আছে, সে কথা কেহই অস্বীকার করিতে পারে না। কিন্তু সেই সুবিধার দ্বারা মনের স্বাভাবিক শক্তিকে একেবারে আচ্ছন্ন করিয়া ফেলিলে বুদ্ধিকে বাবু করিয়া তোলা হয়।

— *রবীন্দ্রনাথ ঠাকুর*

ভারতের একটা mission আছে, একটা গৌরবময় ভবিষ্যৎ আছে, সেই ভবিষ্যৎ ভারতের উত্তরাধিকারী আমরাই। নূতন ভারতের মুক্তির ইতিহাস আমরাই রচনা করছি এবং করব। এই বিশ্বাস আছে বলেই আমরা সব দুঃখ কষ্ট সহ্য করতে পারি, অন্ধকারময় বর্তমানকে অগ্রাহ্য করতে পারি, বাস্তবের নিষ্ঠুর সত্যগুলি আদর্শের কঠিন আঘাতে ধূলিসাৎ করতে পারি।

— *সুভাষচন্দ্র বসু*

Any system of education which ignores Indian conditions, requirements, history and sociology is too unscientific to commend itself to any rational support.

— *Subhas Chandra Bose*

Price : Rs. **400.00**

(*NSOU* -র ছাত্রছাত্রীদের কাছে বিক্রয়ের জন্য নয়)

Published by : Netaji Subhas Open University, DD-26, Sector-1, Salt Lake City, Kolkata-700 064 and
Printed at : Royal Halftone Co., 209/C, Bidhan Saranee, Kolkata-700 006

CBCS • UG • HPH • PHYSICS • CC-PH-10

NETAJI SUBHAS OPEN UNIVERSITY

Choice Based Credit System (CBCS)

SELF LEARNING MATERIAL

HPH

PHYSICS

Mathematical methods in Physics-III

CC-PH-10

Under Graduate Degree Programme

PREFACE

In a bid to standardize higher education in the country, the University Grants Commission (UGC) has introduced Choice Based Credit System (CBCS) based on five types of courses viz. *core, generic, discipline specific, elective, ability and skill enhancement* for graduate students of all programmes at Honours level. This brings in the semester pattern which finds efficacy in sync with credit system, credit transfer, comprehensive continuous assessments and a graded pattern of evaluation. The objective is to offer learners ample flexibility to choose from a wide gamut of courses, as also to provide them lateral mobility between various educational institutions in the country where they can carry their acquired credits. I am happy to note that the University has been recently accredited by National Assessment and Accreditation Council of India (NAAC) with grade "A".

UGC (Open and Distance Learning Programmes and Online Programmes) Regulations, 2020 have mandated compliance with CBCS for U.G. programmes for all the HEIs in this mode. Welcoming this paradigm shift in higher education, Netaji Subhas Open University (NSOU) has resolved to adopt CBCS from the academic session 2021-22 at the Under Graduate Degree Programme level. The present syllabus, framed in the spirit of syllabi recommended by UGC, lays due stress on all aspects envisaged in the curricular framework of the apex body on higher education. It will be imparted to learners over the six semesters of the Programme.

Self-Learning Materials (SLMs) are the mainstay of Student Support Services (SSS) of an Open University. From a logistic point of view, NSOU has embarked upon CBCS presently with SLMs in English / Bengali. Eventually, the English version SLMs will be translated into Bengali too, for the benefit of learners. As always, all of our teaching faculties contributed in this process. In addition to this, we have also requisioned the services of best academics in each domain in preparation of the new SLMs. I am sure they will be of commendable academic support. We look forward to proactive feedback from all stakeholders who will participate in the teaching-learning based on these study materials. It has been a very challenging task well executed by the teaches, officers & staff of the University, and I heartily congratulate all concerned in the preparation of these SLMs.

I wish you all a grand success.

Professor (Dr.) Ranjan Chakrabarti
Vice-Chancellor

Netaji Subhas Open University

Under Graduate Degree Programme
Choice Based Credit System (CBCS)
Subject : Honours in Physics (HPH)
Course : Mathematical methods in Physics-III
Course Code: CC-PH-10

First Print : February, 2023

Printed in accordance with the regulations of the
Distance Education Bureau of the University Grants Commission.

Netaji Subhas Open University

Under Graduate Degree Programme
Choice Based Credit System (CBCS)
Subject : Honours in Physics (HPH)
Course : Mathematical methods in Physics-III
Course Code: CC-PH-10

: Course Writer :
Dr. Md. Abdul Khan
Assistant Professor of Physics,
Aliah University

: Course Editor :
Dr. Gahul Amin
Associate Professor of Physics,
Netaji Subhas Open University

: Format Editor :
Dr. Gahul Amin
Associate Professor of Physics,
Netaji Subhas Open University

Netaji Subhas Open University

UG : Physics (HPH)

Subject : Honours in Physics (HPH)

Course : Mathematical methods in Physics-III

Course Code: CC-PH-10

Unit 1	❒	Complex Analysis	7-70
Unit 2	❒	Integrals Transforms	71-98
Unit 3	❒	Laplace Transforms	99-124
Unit 4	❒	Tensors	125-168

Unit 1 ❑ Complex Analysis

Structure

1.1 Objectives

1.2 Introduction

1.3 Brief Revision of Complex Numbers

1.4 Euler's Formula

1.5 De Moiver's Theorem

1.6 Roots of complex numbers

1.7 Functions of Complex Variables

1.8 Analyticity and Cauchy-Riemann conditions

1.9 Singular functions

1.10 Integration of the function of a complex variable

1.11 Cauchy's integral formula

1.12 Laurent and Taylor's expansion

1.13 Residues and Residue Theorem

1.1 Objective

After studying this section, students will be able to—

- know the historical background of the origin of complex numbers
- understand the basic concept of complex numbers
- understand the basic algebraic properties of complex numbers
- understand the general rules for working with complex numbers
- learn the graphical representation of complex numbers
- understand the uses of complex numbers

1.2 Introduction

The role of complex numbers is inevitable in many branches of physics. Analysis of electric circuits, electromagnetic waves, matter waves, quantum mechanical phenomena, acoustic vibrations, nuclear reactions—all needs the intervention of complex algebra. For instance, we may refer to optics where it appears in the form of complex refractive index causing absorption of electromagnetic waves, in nuclear physics, we introduce complex numbers or complex variables in terms of complex potential to investigate nuclear reactions and many more.

1.3 Brief Revision of Complex Numbers

One of the first mathematicians who realized the need for complex numbers was Italian mathematician Girolamo Cardano (1501-1576). Around 1545, Cardano recognized that his method of solving cubic equations often led to solutions containing the square root of negative numbers. Imaginary numbers did not fully become a part of mathematics, however, until they were studied at length by French-English mathematician Abraham de Moivre (1667-1754), Swiss family of mathematicians named the Bernoullis, Swiss Mathematician Leonhard Euler (1707-1783), and others in the eighteenth century.

Complex Numbers and their Graphical Representation

Complex numbers consist of two parts, one real and one imaginary. An imaginary number is the square root of a negative real number, such as $\sqrt{-4}$. The expression $\sqrt{-4}$ is said to be imaginary because no positive real number can satisfy the condition stated. Meaning, no positive number can be squared to give the value –4. The imaginary number $\sqrt{-1}$ has a special designation i in mathematics.

Complex numbers can be represented as a binomial (a mathematical) expression consisting of one term added to or subtracted from another) of the form $(a+ib)$. In this binomial, a and b represent real numbers and $i=\sqrt{-1}$.

From the above discussions, it is obvious that **complex numbers** can be defined as an ordered pair (x, y) of real numbers that are to be interpreted as points in the **complex plane,** with regular coordinates x and y, just as real numbers x are thought of as points on the real line. When real numbers x are displayed as points $(x, 0)$ on the **real axis**, we write $x = (x, 0)$; and it is quite obvious that the set of complex numbers includes the real numbers as a subset. Complex numbers of the type $(0, y)$ correspond to points on the y-axis and are called **pure imaginary numbers** for non-zero y (i.e. when $y \neq 0$). The y-axis is then referred to as the imaginary axis. It is customary to express a complex number (x, y) by z, so that

$$z = (x, y) \tag{1.1}$$

The real numbers x and y, also respectively called **real** and **imaginary** part of z, and can be donoted as

$$x = \textbf{Re}\ z,\ y = \textbf{Im}\ z \tag{1.2}$$

Two complex numbers z_1 and z_2 will be equal only when they have the same real parts and the same imaginary parts. Thus the statement $z_1 = z_2$ refers to the same point on the complex plane, or on the z plane.

The sum $z_1 + z_2$ and the difference $z_1 - z_2$ of two complex numbers $z_1 = (x_1, y_1)$ and $z_2(x_2, y_2)$ are defined as

$$(x_1, y_1) \pm (x_2, y_2) = (x_1 \pm x_2, y_1 \pm y_2) \tag{1.3}$$

and their product is defined as

$$(x_1, y_1)(x_2, y_2) = (x_1x_2 - y_1y_2, y_1x_2 + x_1y_2) \tag{1.4}$$

It can be noted that the operations defined by means of Eqs (1.3) and (1.4) become the usual operations of addition, substraction and multiplication when restricted to the real numbers:

$$\left.\begin{array}{rcl} (x_1,0) \pm (x_2,0) & = & (x_1 \pm x_2, 0) \\ (x_1,0)(x_2,0) & = & (x_1x_2, 0) \end{array}\right\}$$

Thus, the complex numbers system is an extension of real numbers system. Any complex nmber $z = (x.y)$ can be written as $z = (x,0) + (0,y)$, and it is easy to verify that $(0,1)(y,0) = (0,y)$. Hence $z = (x,0) + (0,1)(y,0)$ and if one thinks of a real number as either x or $(x,0)$ and let i be denote the purely imaginary number (0,1) as depicted in Figure 1.1, then

$$z = x + iy \tag{1.5}$$

Following the convention $z^2 = zz, z^3 = z^2 z, z^4 = z^3 z, etc.$, one may have $i^2 = (0,1)(0,1) = (-1,0)$, or

$$i^2 = -1 \tag{1.6}$$

Since, $(x, y) = x + iy$, definitions (1.3) and (1.4) become

$$(x_1 + iy_1) \pm (x_2 + iy_2) = (x_1 \pm x_2) + i(y_1 \pm y_2) \tag{1.7}$$

and their product is defined as

$$(x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(y_1 x_2 + x_1 y_2) \tag{1.8}$$

Basic Algebraic Properties

Operations with complex numbers mostly follow the same rules as those for real numbers. Two exceptions to those rules arise because of the nature of complex numbers. First, what appears to be an addition operation, $a+bi$, must be left uncombined. Second, the general expression for an imaginary number, such as $i^2 = -1$, violates the rule that the product of two numbers of a like sign is positive. Some of the basic properties of complex numbers follow:

- Closure law if z_1, z_2, z_3 belong to the set S of complex numbers. Then $x_1 + z_2$ and $z_1 z_2$ also belong to the set S.
- The commutative laws

$$z_1 + z_2 = z_2 + z_1, z_1 z_2 = z_2 z_1 \tag{1.9}$$

- The Associative laws

$$(z_1 + z_2) + z_3 = z_2 + z_1, z_1 z_2 = z_2 z_1 \tag{1.10}$$

- The distributive law

$$z(z_1 + z_2) = zz_1 + zz_2 \tag{1.11}$$

- Additive identity 0 = (0,0) and the multiplicative identitity 1=(1,0) for real numbers apply to the entire complex number system. That is,

$$z + 0 = z \, and \, z.1 = z \tag{1.12}$$

for all complex number z.

- For each complex number $z=(x,y)$ there exists an additive inverse of the form

$$-z=(-x,-y), \tag{1.13}$$

which satisfies the equation $z+(-z)=0$. It is to e remembered that there exists only one additive inverse of any given z, since the equation $(x,y)+(u,v)=(0,0)$ implies that $u=-x, v=-y$.

- Multiplicative inverse : For any nonzero complex number $z = (x,y)$, there is a number z^{-1} satisfying the condition $zz^{-1} = 1$. To see it we need real numbers u and v expressed in terms of x and y, such that $(x,y)(u,v) = (1,0) \Rightarrow xu - yv = 1$, $yu + xv = 0$. Solving them one may yet

$$\left.\begin{aligned} u &= \frac{x}{x^2+y^2} \\ v &= \frac{-y}{x^2+y^2} \end{aligned}\right\}$$

Thus the multiplicative inverse of $z = (x, y)$ is

$$z^{-1}=\left(\frac{x}{x^2+y^2},\frac{-y}{x^2+y^2}\right) \quad [z \neq 0] \tag{1.14}$$

When $z=0$, the inverse z^{-1} is undefined, because $z = 0$ means $x^2+y^2=0$, which is not allowed in the Eq. (1.14). Existence of multiplicative inverse help us to prove that if a product z_1z_2 is zero, at least one of the factors z_1 and z_2 must be zero. If $z_1z_2 = 0$ and $z_1 \neq 0$, then the iverse z_1^{-1} exists and any complex number multiplied by zero is zero. So,

$$z_2 = z_2.1 = z_2\left(z_1z_1^{-1}\right) = \left(z_1^{-1}z_1\right)z_2 = z_1^{-1}(z_1z_2) = z_1^{-1}(0) = 0.$$

Thus, if $z_1z_2 = 0$, then either $z_1 = 0$ or $z_2 = 0$; or possibly both z_1 and z_2 are zero. It can also be stated that if two complex numbers z_1 and z_2 are nonzero, then their product z_1z_2 is also nonzero.

- Subtracton and division can be defined in terms of additive and multiplicative inverses:

$$z_1 - z_2 = z_1 + (-z_2) \tag{1.15}$$

$$\frac{z_1}{z_2} = z_1 z_2^{-1} (z_2 \neq 0) \tag{1.16}$$

Using the notations, $z_1 = (x_1, y_1)$ and $z_2 = (x_2, y_2)$, above relations can be expressed as

$$z_1 - z_2 = (x_1, y_1) + (-x_2, -y_2) = (x_1 - x_2, y_1 - y_2) \tag{1.17}$$

$$\left.\begin{aligned} \frac{z_1}{z_2} &= (x_1, y_2)\left(\frac{x_2}{x_2^2 + y_2^2}, \frac{-y_2}{x_2^2 + y_2^2}\right) \\ &= \left(\frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2}, \frac{y_1 x_2 - x_1 y_2}{x_2^2 + y_2^2}\right) \quad [z_2 \neq 0] \end{aligned}\right\} \tag{1.18}$$

Again using the notations, $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, above relations can further be expressed as

$$z_1 - z_2 = (x_1 - x_2) + i(y_1 - y_2) \tag{1.19}$$

$$\frac{z_1}{z_2} = \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + i\frac{y_1 x_2 - x_1 y_2}{x_2^2 + y_2^2} \; [z_2 \neq 0] \tag{1.20}$$

Eq. (1.20) can also be written in an easy to remember form as

$$\frac{z_1}{z_2} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)} \qquad 1.21)$$

- Using Eqs. (1.11) and (1.16), one can also write

$$\left.\begin{aligned} \frac{z_1 + z_2}{z_3} &= (z_1 + z_2) z_3^{-1} \\ &= z_1 z_3^{-1} + z_2 z_3^{-1} \\ &= \frac{z_1}{z_3} + \frac{z_2}{z_3} \end{aligned}\right\} \tag{1.22}$$

- Relation (1.16) with $z_1 = 1$, gives

$$\frac{1}{z_2} = z_2^{-1} \; [z_2 \neq 0] \tag{1.23}$$

- Eq. (1.23) enables us to write Eq. (1.16) in the form

$$\frac{z_1}{z_2} = z_1\left(\frac{1}{z_2}\right) \; [z_2 \neq 0] \tag{1.24}$$

Further one can verify that

$(z_1 z_2)(z_1^{-1} z_2^{-1}) = (z_1 z_1^{-1})(z_2 z_2^{-1}) = 1 \; [z_1 \neq 0,\, z_2 \neq 0]$ and hence that

$z_1^{-1} z_2^{-1} = (z_1 z_2)^{-1}$.

- One can use Eq. (1.24) to verify that

$$\left.\begin{aligned}\left(\frac{1}{z_1}\right)\left(\frac{1}{z_2}\right) &= z_1^{-1} z_2^{-1} \\ &= (z_1 z_2)^{-1} \\ &= \frac{1}{z_1 z_2} \; [z_1 \neq 0, z_2 \neq 0]\end{aligned}\right\} \tag{1.25}$$

- Another useful property which can be derived is

$$\left(\frac{z_1}{z_3}\right)\left(\frac{z_2}{z_4}\right) = \frac{z_1 z_2}{z_3 z_4} \tag{1.26}$$

- The binomial formula involving real numbers remains valid with complex numbers. That is, if z_1 and z_2 are any two nonzero complex numbers, then

$$(z_1 + z_2)^n = \sum_{k=0}^{n} \binom{n}{k} z_1^k z_z^{n-k} \; (n = 1,2,\ldots) \tag{1.27}$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!} (k = 0,1,2,\ldots,n \,\&\, 0! = 1.)$. Applying additive commutation property of complex numbers Eq. ( ) can also be stated in the form

$$(z_1 + z_2)^n = \sum_{k=0}^{n} \binom{n}{k} z_1^{n-k} z_2^k \ (n = 1, 2, \ldots) \tag{1.28}$$

The general rules for working with complex numbers

The following are the basic rules that apply to the algebraic operation of complex numbers.

1. Equality: To be equal, two complex numbers must have equal real parts and equal imaginary parts. That is, assume that we know that the expressions $(a + ib)$ and $(c + id)$ are equal. That condition can be true if and only if $a = c$ and $b = d$.
2. Addition: To add two complex numbers, the real parts and the imaginary parts are added sepatately. The following examples illustrate this rule:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

$$(3 + 5i) + (8 - 7i) = 11 - 2i$$

3. Subtraction: To subtract a complex number, subtract the real part from the real part and the imaginary part from the imaginary part. For example:

$$(a + ib) - (c + id) = (a - c) + i(b - d)$$

$$(6 + 4i) - (3 - 2i) = 3 + i6$$

4. Zero: To equal zero, a complex number must have both its real part and its imaginary part equal to zero. That is, $a + bi = 0$ if and only if $a = 0$ and $b = 0$.
5. Opposites: To form the opposite of a complex number, take the opposite of each part. The opposite of $a + ib$ is $-(a + ib)$, or $-a + i(-b)$. The opposite of $6 - 2i$ is $-6 + 2i$.
6. Multiplication: To form the product of two complex numbers, multiply each part of one number by each part of the other. The product of $(a \times ib) \times (c + id)$ is $ac + iad + ibc + i^2bd$. Since $i^2bd = -bd$, the final product is $ac + adi + bci - bd$. This expression can be expressed as a

complex number as $(ac-bd)+i(ad+bc)$. Similarly, the product $(5-2i)\times(4-3i)$ is 14 – 23*i*.

7. Conjugates: Two numbers whose imaginary parts are opposites are called complex conjugates. The complex numbers $a+bi$ and $a-bi$ are complex conjugates because the terms *bi* have opposite signs. Pairs of complex conjugates have many applications because the product of two complex conjugates is real. For example, $(6-12i)\times(6+12i)=36-144i^2$, or $36+144=180$.
8. Division: Division of complex numbers is restricted by the fact that an imaginary number cannot be divided by itself. Division can be carried out, however, if the divisor is first converted to a real number. To make this conversion, the divisor can be multiplied by its complex conjugate.

Graphical Representation

After complex numbers were discoverd in the eighteenth century, mathematicians searched for possible ways of representing these combinations of real and imaginary numbers. One suggestion was to represent the numbers graphically, as shown in Figure 1.1. In graphical systems, the real part of a complex number is plotted along the horizontal axis (i.e., X-axis) and the imaginary part is plotted on the vertical axis (i.e., Y-axis).

In figure 1.1(a), point P stands for the complex number 3 + 4*i*, point *Q* stands for the complex number –3 + 3*i*, point *R* stands for the complex number –2.5 – 1.5*j* and point *S* stands for the complex number 2 – 2*i*.

In Figure 1.1(b); *P* represents a point in the complex plane to represent the complex number (*x*, *y*) or (*x* + *iy*), such that $x=r\cos\theta, y=r\sin\theta$, $r=\sqrt{x^2+y^2}=|x+iy|$ is called the absolute value or modules of $z=x+iy$ and $\theta=\tan^{-1}\left(\frac{y}{x}\right)$ is called the amplitude or argument of $z=x+iy$ also denote by arg *z*.

Thus, $z=x+iy=r\cos\theta+ir\sin\theta=r(\cos\theta+i\sin\theta)$ represents the polar form of the complex number with polar coordinates *r* and θ.

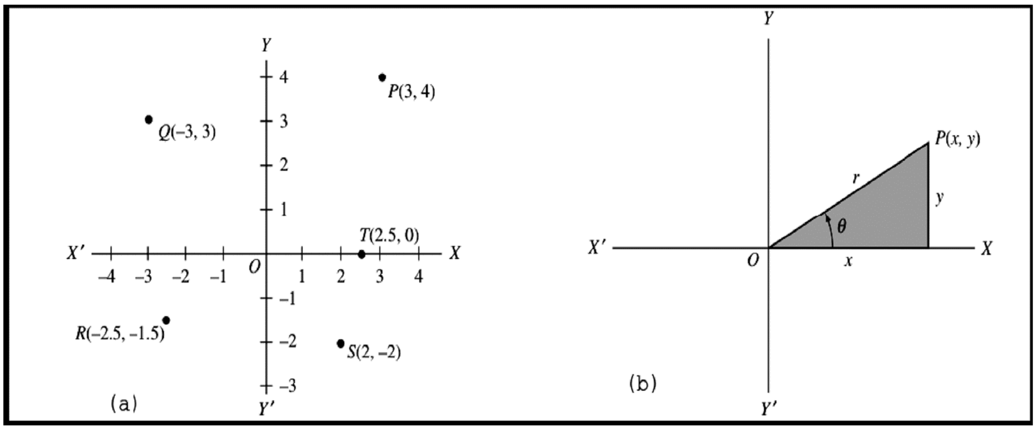


Figure 1.1: Graphical representation of complex numbers: (a) in cartesian coordinates, (b) in polar coordinates

Use of Complex Numbers

For all the ''imaginary'' components they contain, complex numbers occur frequently in scientific and engineering calculations. Whenever the solution to an equation yields the square root of a negative number $\left(\text{such as } \sqrt{-9}\right)$, complex numbers are involved. One of the problems faced by a scientist or engineer, then, is to figure out what the imaginary and complex numbers represent in the real world.

1.4 Euler's Formula

Objectives

After studying this section, students will be able to-

- know the historical background of the Euler's formula
- understand the fundamental relationship between trigonometric functions and complex exponential function
- learn the graphical representation of the complex function
- understand how to interpret Euler's formula
- understand use Euler's formula in defining logarithm of complex numbers
- understand the connection of complex analysis with trigonometry
- learn exponentiation of complex function

Introduction

Euler's formula, named after Leonhard Euler, is a mathematical formula in complex analysis that establishes the fundamental relationship between trigonometric functions and the complex exponential function. Euler's formula states that for any real number x:

$$e^{ix} = \cos x + i\sin x \tag{1.29}$$

which for $x \to -x$ becomes

$$e^{-ix} = \cos x - i\sin x \tag{1.30}$$

where *e* is the base of the natural logarithm, *i* is the imaginary unit, and cos and sin are the trigonometric functions cosine and sine respectively, with the argument *x* given in radians. This complex exponential function is sometimes denoted cis *x* (''cosine plus *i* sine''). The formula is still valid if *x* is a complex number, and so some authors refer to the more general complex version as Euler's formula.[1]

Euler's formula has many applications in mathematics, physics, and engineering. The physicist Richard Feynman called the equation ''our jewel'' and ''the most remarkable formula in mathematics''.[2] When $x = \pi$, Euler's formula evaluates to $e^{i\pi} + 1 = 0$, which is known as Euler's identity.

Interpretation of Euler's Formula

Euler's formula can be interpreted as saying that the function $e^{i\phi}$ is a unit complex number that traces out a unit circle in the complex plane as ϕ ranges through the real numbers. Here ϕ is the angle in radians that a line connecting the origin to a point on the unit circle makes with the positive real axis, in the counterclockwise direction.

The original proof is based on the Taylor series expansions of the exponential function e^z (where *z* is a complex number) and of sin *x* and cos *x* for real numbers *x*. The same proof shows that Euler's formula is even valid for all complex number *x*.

A point in the complex plane can be represented by a complex number written

---

1. Moskowitz, Martin A. (2002). A Course in Complex Analysis in One Variable. World Scientific Publishing Co. p. 7. ISBN 981-02-4780-X.
2. Feynman, Richard P. (1977). The Feynman Lectures on Physics, vol. I. Addison-Wesley. p. 22-10. ISBN 0-201-02010-6.

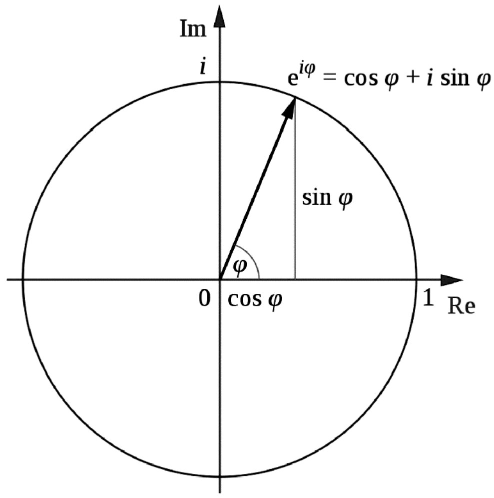


Figure 1.2: Graphical representation of complex number

in cartesian coordinates. Euler's formula provides a means of conversion between cartesian coordinates and polar coordinates. The polar form simplifies the mathematics when used in multiplication or powers of complex numbers. Any complex number $z = x + iy$, and its complex conjugate, $z = x - iy$, can be written as

$$\left.\begin{aligned} x &= x + iy \\ &= |z|(\cos\varphi + i\sin\varphi) \\ &= re^{i\varphi} \\ \bar{z} &= x - iy \\ &= |z|(\cos\varphi - i\sin\varphi) \\ &= re^{-i\varphi} \end{aligned}\right\} \tag{1.31}$$

where $x = \text{Re}(z)$ is the real part of $z, y = \text{Im}(z)$ is the imaginary part of $z, r = |z| = \sqrt{x^2 + y^2}$ is the magnitude of $z, \phi = \arg z = a\tan 2(y, x)$ [ϕ is the argument of z, i.e., the angle between the x axis and the vector z measured counterclockwise in radians, which is defined up to addition of 2π.

Many texts write $\phi = \tan^{-1}\left(\dfrac{y}{x}\right)$ instead $\phi = a\tan 2(y, x)$, but the first equation needs adjustment when $x \leq 0$. This is because for any real x and y not both zero the angles of the vectors (x,y) and $(-x, -y)$ differ by π radians, but have the identical value of tan $\phi = \frac{y}{x}$.

Use of Euler's formula in Defining Logarithm of Complex Numbers

Euler's formula can be used to define the logarithm of a complex number using the definition of the logarithm as the inverse operator of exponentiation:

$$a = e^{\ln a} \tag{1.32}$$

and that

$$e^a e^b = e^{a+b} \tag{1.33}$$

both valid for any complex numbers a and b. Therefore, one can write:

$$z = |z|e^{i\varphi} = e^{\ln|z|}e^{i\varphi} = e^{\ln|z|+i\varphi} \tag{1.34}$$

for any $z \neq 0$. Taking the logarithm of both sides we find

$$\ln z = \ln|z| + i\varphi \tag{1.35}$$

and in fact this can be used as the definition for the complex logarithm. The logarithm of a complex number is thus a multi-valued function, because ϕ is multi-valued.

Finally, the other exponential law

$$\left(e^a\right)^k = e^{ak} \tag{1.36}$$

which can be seen to hold for all integers k, together with Euler's formula, implies several trigonometric identities, as well as de Moivre's formula.

Euler's Formula in Connecting Complex Analysis with Trigonometry

Euler's formula provides a powerful connection between complex analysis and trigonometry, and provides an interpretation of the sine and cosine functions as weighted sums of the exponential function :

$$\left.\begin{aligned}\cos x = \mathrm{Re}\left(e^{ix}\right) = \frac{e^{ix}+e^{-ix}}{2},\\ \sin x = \mathrm{Im}\left(e^{ix}\right) = \frac{e^{ix}-e^{ix}}{2i}\end{aligned}\right\} \tag{1.37}$$

The two equations above can be derived by adding or subtracting

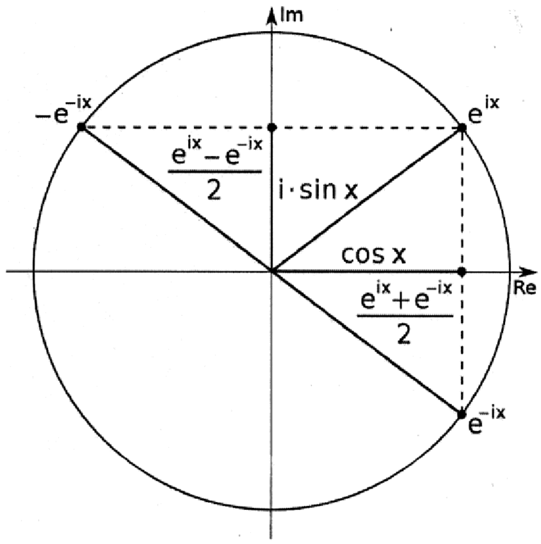


Figure 1.3: Relationship between sine, cosine and exponential function

Euler's formula:

$$\left.\begin{aligned} e^{ix} &= \cos x + i\sin x,\\ e^{-ix} &= \cos(-x)+i\sin(-x) = \cos x - i\sin x\end{aligned}\right\} \tag{1.38}$$

and solving for either cosine or sine. These formulas can even serve as the definition of the trigonometric functions for complex arguments x. For example, letting $x = iy$, we have:

$$\left.\begin{aligned} \cos(iy) &= \frac{e^{-y}+e^{y}}{2} = \cosh(y),\\ \sin(iy) &= \frac{e^{-y}-e^{y}}{2i} = i\left(\frac{e^{y}-e^{-y}}{2}\right) = \sinh(y),\end{aligned}\right\} \tag{1.39}$$

Complex exponentials can simplify trigonometry, because they are easier to manipulate than their sinusoidal components. One technique is simply to convert sinusoids into equivalent expressions in terms of exponentials. After the manipulations, the simplified result is still real-valued. For example:

$$\left.\begin{aligned}
\cos x \cdot \cos y &= \frac{e^{ix}+e^{ix}}{2} \cdot \frac{e^{iy}+e^{-iy}}{2} \\
&= \frac{1}{2} \cdot \frac{e^{i(x+y)}+e^{i(x-y)}+e^{i(-x+y)}+e^{i(-x-y)}}{2} \\
&= \frac{1}{2}\left(\underbrace{\frac{e^{i(x+y)}+e^{-i(x+y)}}{2}}_{\cos(x+y} + \underbrace{\frac{e^{i(x-y)}+e^{-i(x-y)}}{2}}_{\cos(x-y}. \right)
\end{aligned}\right\} \quad (1.40)$$

Another technique is to represent the sinusoids in terms of the real part of a complex expression and perform the manipulations on the complex expression. For example:

$$\left.\begin{aligned}
\cos(nx) &= \mathrm{Re}(e^{inx}) \\
&= \mathrm{Re}\left(e^{i(n-1)x} \cdot e^{ix}\right) \\
&= \mathrm{Re}\left(e^{i(n-1)x} \cdot \left(\underbrace{e^{ix}+e^{-ix}}_{2\cos x} - e^{-ix} \right)\right) \\
&= \mathrm{Re}\left(e^{i(n-1)x} \cdot 2\cos x - e^{i(n-2)x}\right) \\
&= \cos[(n-1)x] \cdot [2\cos(x)] - \cos[(n-2)x]
\end{aligned}\right\} \quad (1.41)$$

This formula is used for recursive generation of cos nx for integer values of n and arbitrary x (in radians).

Exponentiation of Complex Function

The exponential function e^x for real values of x may be defined in a few different equivalent ways (see Characterizations of the exponential function). Several of these methods may be directly extended to give definitions of e^z for complex values of z simply by substituting z in place of x and using the complex algebraic operations. In particular, we may use either of the two following definitions, which are equivalent. From a more advanced perspective, each of these definitions may be interpreted as giving the unique analytic continuation of e^x to the complex plane.

Power Series Definition

For complex z

$$\left.\begin{aligned} e^z &= 1+\frac{z}{1!}+\frac{z^2}{2!}+\frac{z^3}{3!}+\ldots \\ &= \sum_{n=0}^{\infty}\frac{z^n}{n!} \\ e^z &= 1+\frac{z}{1!}+\frac{z^2}{2!}+\frac{z^3}{3!}+\ldots \\ &= \sum_{n=0}^{\infty}\frac{z^n}{n!} \end{aligned}\right\} \tag{1.42}$$

Using the ratio test, it is possible to show that this power series has an infinite radius of convergence and so defines e^z for all complex z.

Limit definition

For complex z

$$e^z = \lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n \tag{1.43}$$

Here, n is restricted to positive integers, so there is no question about what the power with exponent n means.

Proofs: Various proofs of the formula are possible.

Using power series

Here is a proof of Euler's formula using power-series expansions, as well as basic facts about the powers of i:

$$\left.\begin{array}{llll} i^0 = 1, & i^1 = i, & i^2 = -1, & i^3 = -i, \\ i^4 = 1, & i^5 = i, & i^6 = -1, & i^7 = -i, \\ \vdots & \vdots & \vdots & \vdots \end{array}\right\} \tag{1.44}$$

Using now the power-series definition from above, we see that for real values of x

$$\left.\begin{aligned} e^{ix} &= 1+ix+\frac{(ix)^2}{2!}+\frac{(ix)^3}{3!}+\frac{(ix)^4}{4!}+\frac{(ix)^5}{5!}+\frac{(ix)^6}{6!}+\frac{(ix)^7}{7!}+\frac{(ix)^8}{8!}+\ldots \\ &= 1+ix-\frac{x^2}{2!}-\frac{ix^3}{3!}+\frac{x^4}{4!}+\frac{ix^5}{5!}-\frac{x^6}{6!}-\frac{ix^7}{7!}+\frac{x^8}{8!}+\ldots \\ &= \left(1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\frac{x^8}{8!}-\ldots\right)+i\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots\right) \\ &= \cos x+i\sin x. \end{aligned}\right\} \qquad (1.45)$$

In the last step, we have simply recognized the Maclaurin series for cos x and sin x. The rearrangement of terms is justified because each sereies is convergent.

Using polar coordinates :

Another proof is based on the fact that all complex numbers can be expressed in polar coordinates. Therefore, for some r and θ depending on x,

$$e^{ix} = r(\cos\theta + i\sin\theta) \qquad (1.46)$$

No assumptions are being made about r and θ; they will be determined in the course of the proof. From any of the definitions of the exponential function it can be shown that the derivative of e^{ix} is ie^{ix}. Therefore, differentiating both sides gives

$$ie^{ix} = (\cos\theta + i\sin\theta)\frac{dr}{dx} + r(-\sin\theta + i\cos\theta)\frac{d\theta}{dx}. \qquad (1.47)$$

Substituting $r(\cos\theta + i\sin\theta)$ for e^{ix} and equating real and imaginary parts in this formula gives

$$\frac{dr}{dx} = 0$$

and

$$\frac{d\theta}{dx} = 1$$

Thus, r is a constant, and θ is $x+C$ for some constant C. The initial values $r(0) = 1$

and $\theta(0) = 0$ come from $e^{0i} = 1$, giving $r = 1$ and $\theta = x$. This proves the formula

$$e^{ix} = 1(\cos x + i \sin x) = \cos x + i \sin x. \quad (1.48)$$

1.5 De Moivre's theorem

Objectives

After studying this section, students will be able to-

- understand the basic concepts of de Moivre's theorem
- use de Moivre's Theorem to calculate powers of complex numbers
- derive Euler's formula from de Moivre's theorem
- solve problems on complex numbers based on de Moivre's theorem

Introduction

Abraham de Moivre introduced a very important formula named after him that connects complex numbers to trigonometry. De Miovre's formula also known as de Moivre's theorem or de Moivre's identity in mathematics, states that for any real number x and integer n it holds that

$$(\cos(x) + i \sin(x))^n = \cos(nx) + i \sin(nx) \quad (1.49)$$

where i is the imaginary unit ($i^2 = -1$). By expanding the left-hand side and then comparing the real and imaginary parts under the assumption that x is real, it is possible to derive useful expressions for $\cos(nx)$ and $\sin(nx)$ in terms of $\cos(x)$ and $\sin(x)$. The formula is valid for only integer n. A generalization of this formula is valid for other exponents and those can be used to find explicit expressions for the n-th roots of unity, using the relation $z^n = 1$ for complex z. If we have any arbitrary complex number z, we can write it in the complex polar form as

$$z = (r \cos\theta) + i(r \sin\theta) \quad (1.50)$$

where r, θ are real and $i^2 = -1$. Furthermore, if we have another complex number

$$v = (\rho \cos\phi) + i(\rho \sin\phi) \quad (1.51)$$

Then the product of z and v is given by

$$\begin{aligned} zv &= \left[(r\cos\theta) + i(r\sin\theta)\right] \times \left[(\rho\cos\phi) + i(\rho\sin\phi)\right] \\ &= r\rho\left[\cos(\theta+\phi) + i\sin(\theta+\phi)\right] \end{aligned} \quad (1.52)$$

In other words, the moduli of the complex number are multiplied together while the arguments are added together. The above formula can be extended to any integer n.

Generalization of de Moivre's formula

If $z_1 = x_1 + iy_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = x_2 + iy_2 = r_2(\cos\theta_2 + i\sin\theta_2)$, then one can show that

$$z_1 z_2 = r_1 r_2\left[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)\right] \quad (1.53)$$

and

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}\left[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 = \theta_2)\right] \quad (1.54)$$

In general,

$$z_1 z_2 ... z_n = r_1 r_2 .. r_n \left\{\cos(\theta_1 + \theta_2 + ... + \theta_n) + i\sin(\theta_1 + \theta_2 + .. + \theta_n)\right\} \quad (1.55)$$

If, $z_1 = z_2 = z_3 = ... = z_n = z$, $r_1 = r_2 = r_3 = ... = r_n = r$ and

$\theta_1 = \theta_2 = \theta_3 = ... = \theta_n = \theta$, then from (1.55) we have

$$z^n = r^n\left\{\cos(n\theta) + i\sin(n\theta)\right\} \quad (1.56)$$

$$r^n\left\{\cos(\theta) + i\sin(\theta)\right\}^n = r^n\left\{\cos(n\theta) + i\sin(n\theta)\right\} \quad (1.57)$$

$$\left\{\cos(\theta) + i\sin(\theta)\right\}^n = \left\{\cos(n\theta) + i\sin(n\theta)\right\} \quad (1.58)$$

Eq. (1.58) is the De moivre's formula.

Derivation of Euler's formula from de Moivre's theorem

De Moivre's theorem can easily be derived from Euler's formula

$$e^{ix} = \cos x + i\sin x \quad (1.59)$$

and the exponential law for integer powers

$$\left(e^{ix}\right)^n = e^{inx}. \quad (1.60)$$

Then, by Euler's formula,

$$e^{inx} = \cos nx + i\sin nx. \tag{1.61}$$

Proof of de Moiver's theorem by induction (for integer *n*)

The truth of De Moivre's theorem can be established by using mathematical induction for natural numbers and extended to all integers from there. For an integer n, call the following statement $S(n)$:

$$(\cos x + i\sin x)^n = \cos nx + i\sin nx. \tag{1.62}$$

For $n > 0$, we proceed by mathematical induction. $S(1)$ is clearly true. For our hypothesis, we assume $S(k)$ is true for some natural k. That is we assume

$$(\cos x + i\sin x)^n = \cos kx + i\sin kx. \tag{1.63}$$

Now, considering $S(k + 1)$:

$$\left.\begin{aligned}(\cos x + i\sin x)^{k+1} &= (\cos x + i\sin x)^k(\cos x + i\sin x)\\ &= (\cos kx + i\sin kx)(\cos x + i\sin x)\\ &\quad \text{[by the induction hypothesis]}\\ &= \cos(kx)\cos x - \sin(kx)\sin x\\ &\quad +i(\cos(kx)\sin x + \sin(kx)\cos x\\ &= \cos\left((k+1)x\right) + i\sin\left((k+1)x\right)\\ &\quad \text{[by the trigonometric identities]}\end{aligned}\right\} \tag{1.64}$$

We deduce that $S(k)$ implies $S(k+1)$. By the principle of mathematical induction it follows that te result is true for all natural numbers. Now $S(0)$ is clearly true since $\cos(0x) + i\sin(0x) = 1 + 0i = 1$. Finally, for the negative integer cases, we consider an exponent of $-n$ for natural n.

$$\left.\begin{aligned}(\cos x + i\sin x)^{-n} &= \left((\cos x + i\sin x)^n\right)^{-1}\\ &= (\cos nx + i\sin nx)^{-1}\\ &= \cos(-nx) + i\sin(-nx).\end{aligned}\right\} \tag{1.65}$$

where the last equation is the result of the identity

$$z^{-1} = \frac{\bar{z}}{|z|^2}, \tag{1.66}$$

for $z = \cos(nx) + i\sin(nx)$. Hence, $S(n)$ holds for all integers n.

Formulae for cosine and sine

Using the binomial expansion formula

$$(\cos x + i\sin x)^n = \sum_k^n \left({}^nC_k\right)(\cos x)^{n-k}(i\sin x)^k \tag{1.67}$$

one can extract the formula for cosine and sine, where ${}^nC_k$ are the usual binomial coefficients defined as ${}^nC_k = \frac{n!}{(n-k)!k!}$. Equality of two complex numbers, demands equality of both of the real parts and of the imaginary parts of both members of the equation. If x, cos x and sin x are real numbers, then the identity of thee parts can be expressed in terms of binomial coefficients uggested by 16th century French mathematician François Viète:

$$\left.\begin{aligned} \sin(nx) &= \sum_{k=0}^n {}^nC_k(\cos x)^k(\sin x)^{n-k}\sin\frac{(n-k)\pi}{2} \\ \cos(nx) &= \sum_{k=0}^n {}^nC_k(\cos x)^k(\sin x)^{n-k}\cos\frac{(n-k)\pi}{2} \end{aligned}\right\} \tag{1.68}$$

Concrete instancess of these equations for n = 2 and n = 3 follows:

$$\left.\begin{aligned} \cos 2x &= (\cos x)^2 + \left((\cos x)^2 - 1\right) &&= 2\cos^2 x - 1 \\ \sin 2x &= 2(\sin x)(\cos x) &&= 2\sin x\cos x \\ \cos 3x &= (\cos x)^3 + 3\cos x\left((\cos x)^2 - 1\right) &&= 4\cos^3 x - 3\cos x \\ \sin 3x &= 3(\cos x)^2(\sin x) - (\sin x)^3 &&= 3\sin x - 4\sin^3 x. \end{aligned}\right\} \tag{1.69}$$

1.6 Roots of complex numbers

Objectives

After studying this section, students will be able to-

- apply de Moivre's theorem to find roots of a complex number
- understand the geometrical interpretation of the roots of a complex number
- find n-th roots of unity
- understand how to get solution of polynomial equations

Introduction

A number u is said to be an n-th root of complex number z, if u^n = z or $u = z^{1/n}$.

Theorem 1.6.1 *Every complex number has exactly n distinct n-th roots.*

$$\left.\begin{aligned} Let, z &= r(\cos\theta + i\sin\theta); \\ and, u &= \rho(\cos\alpha + i\sin\alpha) \end{aligned}\right\}$$

$$\left.\begin{aligned}
Then,\ r(\cos\theta + i\sin\theta) &= [\rho(\cos\alpha + i\sin\alpha)]^n \\
&= \rho^n(\cos\alpha + i\sin\alpha)^n \\
&= \rho^n(\cos n\alpha + i\sin n\alpha) \\
&\quad \text{[By deMoivre's theorem]} \\
&\Rightarrow \rho^n = r, n\alpha = \theta + 2\pi k \\
&\quad [where,\ k = 0, 1, 2, \ldots\ldots, (n-1) \\
or\quad & \rho = r^{1/n}, \alpha = \frac{\theta}{n} + \frac{2\pi k}{k} \\
\text{Hence, } u &= z^{1/n} \\
&= r^{1/n}\left[\cos\left(\frac{\theta}{n} + \frac{2\pi k}{n}\right) + i\sin\left(\frac{\theta}{n} + \frac{2\pi k}{n}\right)\right]
\end{aligned}\right\}$$

Geometrical interpretation of the roots of a complex number

For a point $z = re^{i\theta}$, lying on a circle of radius r and centred at the origin, as θ increases, the point z moves around the circle in the anticlockwise direction. And in particular when θ is increased by 2π, we return to the starting point, and the same is true if θ is increased further by 2π. This means that two non-vanishing complex numbers $z_2 = r_1$ and $\theta_2 = \theta_1 + 2\pi k$ with $\theta_2 = \theta_1 + 2\pi k$ integer.

Thus the expression $z = re^{i\theta}$ raised to any integrer power n may be used to find the n-th root of the nonozero complex number $z_0 = r_0 e^{i\theta_0}$ if $z^n = z_0$ or $r^n e^{in\theta} = r_0 e^{i\theta_0}$.

Which gives: $z = z_0^{1/n} = r_0^{1/n} e^{i\left(\frac{\theta_0}{n} + \frac{2\pi k}{n}\right)}$ where $k = 0,1,2,3,\ldots,(n-1)$. In general the k^{th} root of the non-zero complex number $z_0 = r_0 e^{i\theta_0}$ can be expressed as:

$$c_k = r_0^{\frac{1}{n}} e^{i\left(\frac{\theta_0}{n} + \frac{2\pi k}{n}\right)}, \left[k = 0,1,2,\ldots,(n-1)\right] \tag{1.70}$$

The n^{th} roots of unity

The solutions of the equation $z^n = 1$ where n is a positive integer are called the n^{th} roots of unity and are given by

$$z = \cos\left(\frac{2k\pi}{n}\right) + i\sin\left(\frac{2k\pi}{n}\right) = e^{\frac{i2k\pi}{n}} \quad \left[k = 0,1,2,\ldots,(n-1)\right] \tag{1.71}$$

If we assume $\omega = \cos\left(\frac{2\pi}{n}\right) + i\sin\left(\frac{2\pi}{n}\right) = e^{\frac{2\pi i}{n}}$, then n roots will be 1, $\omega, \omega^2, \omega^3, \ldots, \omega^{n-1}$ respectively. Geometrically, the roots represent the n vertices of an n-sided regular polygon inscribed in a circle of radius unity with centre at the origin. This circle has the equation $|z| = 1$ and is often called the unit circle.

Polynomial Equations

More often we require solutions of polynomial equations of the form

$$c_0 z^n + c_1 z^{n-1} + c_2 z^{n-2} + \ldots\ldots + c_r z^{n-r} + \ldots + c_n = 0 \tag{1.72}$$

where $c_0 \neq 0, c_1, \ldots, c_n$ are given complex numbers and n is a positive integer called the degree of the equation. Such solutions are also called zeros of the polynomial on the left of Eq. (1.72) or roots of the equation. A very important theorem called the fundamental theorem of algebra states that every polynomial equation of the form (1.72) has at least one root in the complex plane. From this, we can show that it has, in fact, n complex roots, some or all of which may be identical. If $z_1, z_2, \ldots, z_n$ are n roots of the polynomial (1.72), then we can write.

$$c_0(z-z_1)(z-z_2)(z-z_3)(z-z_4)\ldots\ldots(z-z_n)=0 \qquad (1.73)$$

which is called the factored form of the polynomial equation.

Solved examples on basic operations of complex numbers

1. $(3+2i)+(-6-3i)=?$

   Solution :

   $(3+2i)+(-6-3i)=(3-6)+i(2-3)=-3-i$

2. $(-7-3i)+(5+4i)=?$

   Solution :

   $(-7-3i)+(5+4i)=(-7+5)+i(-3+4)=-2+i$

   Example 1&2 illustrate commutative law of addition.

3. $(6-6i)-(2i-7)=?$

   Solution :

   $(8-6i)-(2i-7)=8-6i-2i+7=15=8i$

4. $(5+3i)+\{(-1+2i)+(7-5i)\}=(5+3i)+\{-1+2i+7-5i\}=$

   $(5+3i)+(6-3i)=11$

   Examples 3 & 4 illustrate the associative law of addition.

5. $(2-3i)(4+2i)=?$

   Solution :

   $(2-3i)(4+2i)=2(4+2i)-3i(4+2i)=8+4i-12i-6i^2=$

   $8+4i-12i+6=14-8i$

6. $(4+2i)(2-3i)=?$

Solution :

$(4+2i)(2-3i) = 4(2-3i)+2i(2-3i) = 8-12i+4i-6i^2 =$

$8-12i+4i+6 = 14-8i$

Examples 5 & 6 illustrate the commutative law of multiplication.

7. $(2-i)(-3+2i)(5-4i) = ?$

Solution :

$(2-i)(-3+2i)(5-4i) = (2-i)\{-15+12i+10i-8i^2\} =$

$(2-i)(-7+22i) = -14+44i+7i-22i^2 = 8+51i$

8. $(2-i)(-3+2i)(5-4i) = ?$

Solution :

$(2-i)(-3+2i)(5-4i) = \{-6+4i+3i-2i^2\}(5-4i) = \{-6+4i+$

$3i+2\}(5-4i) = (-4+7i)(5-4i) = \{-20+16i+35i-28i^2\} =$

$\{-20+16i+35i+28\} = 8+51i$

Examples 7 & 8 illustrate the associative law of multiplication.

9. $(-1+2i)\{(7-5i)+(-3+4i)\} = ?$

Solution :

$(-1+2i)\{7-5i)+(-3+4i)\} = (-1+2i)(4-i) = -4+i+8i-2i^2 =$

$-4+i+8i+2 = -2+9i$

Alter :

$(-1+2i)\{(7-5i)+(-3+4i) = (-1+2i)(7-5i)+(-1+2i)(-3+4i) =$

$\{-7+5i+4i-10i^2\}+\{3-4i-6i+8i^2\} = (3+19i)+(-5-10i) = -2+9i$

This example illustrates the distributive law.

10. $\frac{3-2i}{-1-i} = ?$

Solution :

$$\frac{3-2i}{-1-i} = \left(\frac{3-2i}{-1-i}\right)\left(\frac{-1+i}{-1+i}\right) = \left(\frac{-3+3i+2i-2i^3}{(-1)^2-(i)^2}\right) = \left(\frac{-1+5i}{1+1}\right) = -\frac{1}{2}+\frac{5}{2}i$$

Alter : If $\frac{3-2i}{-1-i} = (p+qi)$ (say),

then, $3-2i = (-1-i)(p+qi) = -p-qi-pi-qi^3 = -p-i(p+q)+q =$

$(-p+q)-i(p-q) \Rightarrow -p+q = 3 \,\&\, p+q = 2 \Rightarrow p =$
$-\frac{1}{2} \,\&\, q = \frac{5}{2} \Rightarrow p+qi = -\frac{1}{2}+\frac{5}{2}i$

11. $\frac{3i^{30}-i^9}{2i-1} = ?$

Solution :

$$\frac{3i^{30}-i^{19}}{2i-1} = \frac{3(i^2)^{15}-(i^2)^9 i}{2i-1} = \frac{3(-1)^{15}-(-1)^9 i}{2i-1} = \frac{-3+1}{2i-1} = \frac{(-3+i)(-1-2i)}{(-1+2i)(-1-2i)}$$

$$\frac{3+6i-i+2}{(-1)^2-(2i)^2} = \frac{5+5i}{1+4} = 1+i$$

12. If $z_1 = 2+i, z_2 = 3-2i, z_3 = -\frac{1}{2} = \frac{\sqrt{3}}{2}i$, then evaluate (a) $|3z_1-4z_2|$,

(b) $z_1^3-3z_1^2+4z_1-8$, (c) $(\bar{z}_3)^4$, (d) $\left|\frac{2z_2+z_1-i}{2z_1-z_2+3-i}\right|$.

Solution :

(a) $|3z_1-4z_2| = |3(2+i)-4(3-2i)| = |6+3i-12+8i| = |-6+11i|$
$= \sqrt{(-6)^2+(11)^2} = \sqrt{36+121} = \sqrt{157}$

(b) $z_1^3 - 3z_1^2 + 4z_1 - 8 = (2+i)^3 - 3(2+i)^2 + 4(2+i) - 8 =$
$\{(2)^3 + 3(2)^2(i) + 3(2)(i)^2 + i^3\} - 3(4 + 4i + i^2) + 8 + 4i - 8 =$
$8 + 12i - 6 - i - 12 - 12i + 3 + 8 + 4i - 8 = -7 + 3i$

(c) $(\bar{z}_3)^4 = \left(\overline{-\frac{1}{2} + \frac{\sqrt{3}}{2}i}\right)^4 = \left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right)^4 = \left[\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right)^2\right]^2$

$\left(\frac{1}{4} + \frac{\sqrt{3}}{2}i - \frac{3}{4}\right)^2 = \left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right)^2 = \left(\frac{1}{4} - \frac{\sqrt{3}}{2}i - \frac{3}{4}\right) = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$

(d) $\left|\frac{2z_2 + z_1 - 5 - i}{2z_1 - z_2 + 3 + i}\right|^2 = \left|\frac{2(3-2i) + (2+i) - 5 - i}{2(2+i) - (3-2i) + 3 - i}\right|^2 = \left|\frac{6 - 4i + 2 + i - 5 - i}{4 + 2i - 3 + 2i + 3 - i}\right| =$

$\left|\frac{3-4i}{4+31}\right|^2 = \frac{|3-4i|^2}{|4+3i|^2} = \frac{\left(\sqrt{3^2+4^2}\right)^2}{\left(\sqrt{4^2+3^2}\right)^2} = \frac{25}{25} = 1$

13. Show that : (a) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$. (b) $|z_1 z_2| \le |z_1||z_2|$.

Solution :

Let $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$. Then

(a) $\overline{z_1 + z_2} = \overline{(x_1 + iy_1) + (x_2 + iy_2)} = \overline{(x_1 + x_2) + i(y_1 + y_2)} =$
$(x_1 + x_2) - i(y_1 + y_2) = (x_1 - iy_1) + (x_2 - iy_2) = \overline{x_1 + iy_1} + \overline{x_2 + iy_2} =$
$\bar{z}_1 + \bar{z}_2$

(b) $|z_1 z_2| = |(x_1 + iy_1)(x_2 + iy_2)| = |(x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2)|$
$= \sqrt{x_1^2x_2^2 + y_1^2y_2^2 - 2x_1x_2y_1y_2 + x_1^2y_2^2 + 2x_1x_2y_1y_2}$
$= \sqrt{x_1^2x_2^2 + y_1^2y_2^2 + x_1^2y_2^2 + y_1^2x_2^2} = \sqrt{\left(x_1^2 + y_1^2\right)\left(x_2^2 + y_2^2\right)}$
$= \sqrt{x_1^2 + y_1^2}\,\sqrt{x_2^2 + y_2^2} = |x_1||x_2|$

Alter :

$$|z_1z_2|^2 = (z_1z_2)(\overline{z_1z_2}) = z_1z_2\bar{z}_1\bar{z}_2 = (z_1\bar{z}_1)(z_2\bar{z}_2) = |z_1|^2|z_2|^3$$
$$\Rightarrow |z_1z_2| = |z_1||z_2|$$

14. Prove (a) $|z_1+z_2| \le |z_1|+|z_2|$, (b) $\|z_1+z_2+z_3| \le |z_1|+|z_2|+|z_3|$,

(c) $|z_1-z_2| \ge |z_1|-|z_2|$

Solution :

Let $z_1 = x_1+iy_1, z_2 = x_2+iy_2$. Then

(a) $|z_1 = z_2| \le |z_1|+|z_2| \Rightarrow |(x_1+iy_1)+(x_2+iy_2)| \le \sqrt{x_1^2+y_1^2} +$
$\sqrt{x_2^2+y_2^2} \Rightarrow \sqrt{(x_1+x_2)^2+(y_1+y_2)^2} \le \sqrt{x_1^2+y_1^2}+\sqrt{x_2^2+y_2^2}$

Squaring both sides we have

$$(x_1+x_2)^2+(y_1+y_2)^2 \le (x_1^2+y_1^2)+(x_2^2+y_2^2)+2\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2}$$
$$\Rightarrow x_1x_2+y_1y_2 \le \sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2}$$

Squaring both sides once agin we have

$$(x_1x_2+y_1y_2)^2 \le (x_1^2+y_1^2)(x_2^2+y_2^2) \Rightarrow 2x_1x_2y_1y_2 \le (x_1^2y_2^2+y_1^2x_2^2)$$
$$\Rightarrow 0 \le (x_1y_2-y_1x_2)^2, \text{ or, } (x_1y_2-y_1x_2)^2 \ge 0$$

[which is true, hence the given relation is correct]

(b) $|(z_1+z_2)+z_3| \le |z_1+z_2|+|z_3| \le |z_1|+|z_2|+|z_3|$

(c) $|z_1| = |z_1-z_2+z_2| \le |z_1-z_2|+|z_2| \Rightarrow |z_1-z_2| \ge |z_1|-|z_2|$.

Replacing z_2 by $-z_2$ we also have $|z_1+z_2| \ge |z_1|-|z_2|$.

15. Find all roots of the number - 16.

Solution :

The given number can be expressed in the standard form as:

$$z_0 = r_0e^{i\theta_0} = 16e^{i\pi}. \text{ So, } r^n = r_0 = 16 = 2^4 \Rightarrow r = 2,\ n = 4,\ \theta_0 = \pi.$$

Hence, $\theta = \frac{\theta_0}{n} + \frac{2\pi k}{n} = \frac{\pi}{4} + \frac{2\pi k}{4}$. Thus, the given number has four roots (c_k) corresponding to k = 0, 1, 2, and 3 which are : $c_0 = 2e^{i\frac{\pi}{4}} = 2\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) = 2\left(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right) = \sqrt{2}(1+i)$, Similarly, $c_1 = \sqrt{2}(-1+i)$, $c_2 = \sqrt{2}(-1-i)$, $c_3 = \sqrt{2}(1-i)$.

16. Prove that $\cot^{-1} z = \frac{1}{2i}\log\left(\frac{z+i}{z-i}\right)$

Solution :

Let $w = \cot^{-1} z$. Then, $z = \cot w$. Now, $\frac{z+i}{z-i} = \frac{\cot w + i}{\cot w - i} = \frac{\cos w + i\sin w}{\cos w - i\sin w} = \frac{e^{iw}}{e^{-iw}} = e^{2iw}$. Taking logarithm on both sides, we get: $\log\left(\frac{z+i}{z-i}\right) = 2iw = 2i\cot^{-1} z \Rightarrow \cot^{-1} z = \frac{1}{2i}\log\left(\frac{z+i}{z-i}\right)$ (Proved)

17. Prove that $\sec^{-1} z = \frac{1}{i}\log\left(\frac{1+\sqrt{z^2-1}}{z}\right)$

Solution :

Let $w = \sec^{-1} z$. Then, $z = \sec w$. Now, $\frac{1+\sqrt{1-z^2}}{z} = \frac{1+\sqrt{1-\sec^2 w}}{\sec w} = \frac{1+\sqrt{-\tan^2 w}}{\sec w} = \frac{1+i\tan w}{\sec w} = \frac{\frac{\cos w + i\sin w}{\cos w}}{\frac{1}{\cos w}} = \cos w + i\sin w = e^{iw}$. Taking logarithm on both sides, we get: $\log\left(\frac{1+\sqrt{1-z^2}}{z}\right) = iw = i\sec^{-1} z \Rightarrow \sec^{-1} z = \frac{1}{i}\log\left(\frac{1+\sqrt{1-z^2}}{z}\right)$ (Proved).

18. Solve the quadratic equation $az^2 + bz + c = 0,\ a \neq 0.$

Solution :

$$az^2 + bz + c = 0$$

$$\text{or,}\quad z^2 + \frac{b}{a}z + \frac{c}{a} = 0$$

$$\text{or,}\quad z^2 + 2(z)\left(\frac{b}{2a}\right) + \left(\frac{b}{2a}\right)^2 + \frac{c}{a} = \left(\frac{b}{2a}\right)^2$$

$$\text{or,}\quad \left(z + \frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2 + \frac{c}{a} = 0$$

$$\text{or,}\quad \left(z + \frac{b}{2a}\right)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a}$$

$$\text{or,}\quad \left(z + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

$$\text{or,}\quad z + \frac{b}{2a} = \pm\frac{b^2 - 4ac}{4a^2}$$

$$\text{or,}\quad z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

19. Solve the equation $z^2 + (2i - 3)z + 5 - i = 0.$

Solution :

Comparing the given equation with the standard quadratic equation :

$az^2 + bz + c = 0$, we have, $a = 1, b = 2i - 3, c = 5 - i$ Thus

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$= \frac{-(2i-3) \pm \sqrt{(2i-3)^3 - 4(1)(5-i)}}{2(1)}$$

$$= \frac{3 - 2i \pm \sqrt{-15 - 8i}}{2}$$

$$= \frac{3-2i \pm \sqrt{1-8i+16i^2}}{2}$$

the solutions are : $= \dfrac{3-2i \pm \sqrt{1-8i+16i^2}}{2}$

$$= \frac{3-2i \pm \sqrt{(1-4i)^2}}{2}$$

$$= \frac{3-2i \pm (1-4i)}{2}$$

$$= \left[\frac{3-2i+(1-4i)}{2}\right] \text{or} \left[\frac{3-2i-(1-4i)}{2}\right]$$

$$= (2-3i) \text{ or } (1+i).$$

1.7 Functions of Complex Variables

Objectives

After studying this section, students will be able to-

- understnad the basic structure of the function of a complex variable
- understand mathematical representation of complex variables
- understand concepts of single and multiple-valued functions
- concept of branch
- learn concepts of limit, continuity, and differentiability of functions of complex variables
- learn to find inverse function corresponding to a function of complex variables

Introduction

The theory of functions of a complex variable, also known as complex variables or complex analysis, is one of the beautiful as well as useful branches of mathematics. Although originating in an atmosphere of mystery, suspicion, and distrust, as evidenced by the terms imaginary and complex present in the literature, it was finally placed on a sound foundation in the 19th century through the efforts of Cauchy, Riemann, Weierstrass, Gauss, and other great mathematicians." It has many applications in branches of mathematics including algebraic geometry, number theory, analytic combinatorics, applied mathematics, etc. In physics, it appears in hydrodynamics, thermodynamics, and particularly in quantum mechanics. Complex analysis has its uses in engineering disciplines such as nuclear, aerospace, mechanical, and electrical engineering. In the field of engineering and science, many complicated integrals of real functions are solved with the help of functions of a complex variable.

Functions of complex variables

An arbitrary set of complex numbers represented by the symbol $z(= x + iy)$ which is made up of a real component (x) and an imaginary component (y), is called a complex variable. Now if, to each value of that complex variable z, there corresponds one or more values of another complex variable say $w(= u+iv)$, then we say that w is a function of z and express the relation as $w = f(z)$, where the variable z is called independent variable while w is called dependent variable. The value of the function $f(z)$ at some point say at $z = a$ is often written $f(a)$.

Thus we see that $f(z)$ is a function of complex variable z and is denoted by w. $w = f(z)$ and $w = u + iv$, where $u = u(x, y)$ and $v = v(x, y)$ are the real and imaginary components of $f(z)$.

Mathematical representation of complex variables

- Cartesian form : $z = x + iy, i = \sqrt{-1}$
- Polar form : $z = r(\cos\theta + i\sin\theta)$
- Exponential form : $z = re^{1\theta}$

Examples on mathematical representation of complex variables

1. If $f(z) = z^3$, then $f(2i) = (2i)^3 = -8i$.

2. If $f(z) = \frac{1}{z}$, $\left[f(z) = u + iv,\ z = x + iy,\ u = \dfrac{x}{x^2+y^2},\ v = \dfrac{-y}{x^2+y^2}\right]$

3. $f(z) = z\bar{z}\left[f(z) = u + iv,\ z = x + iy,\ u = x^2 + y^2,\ v = 0\right]$

4. $f(z) = \dfrac{z-i}{z+1}$

Single valued function

For $w = f(z)$, if each value of z corresponds to only one value of w, then we may say that w is a single-valued function of z or that $f(z)$ is single-valued.

Multiple-valued function

For $w = f(z)$, when each value of z correspods to more than one value of w, then we say that w is a multiple-valued or many-valued function of z.

Example of single and multiple-valued functions

1. If $w = z^3$, then to each value of z there is only one value of w. Hence, $w = f(z) = z^3$ is a single-valued function of z.
2. If $w^2 = z$, then to each value of z there are two values of w. Hence, $w^2 = z$ defines a multiple-valued (here two-valued) function of z.

Branch

A multiple-valued function can be considered as a collection of singlevalued functions, each member of which is called a branch of the function. /Usually, one of the branches of the multiple-valued function is designated as a principal branch and the value of the function corresponding to this branch as the principal value.

Limit of a funtion of a complex variable

Let $f(z)$ be a single-valued function defined at all points in the neighbourhood of the point $z0$. Then the limit of $f(z)$ as z approaches z_0 is w_0 expressed as: $\lim_{z \to z_0} f(z) = w_0$

Continuity of a funtion of complex variable

The function $f(z)$ of a complex variable z is said to be continuous at the point z_0 if for any given positive number $\in$, we can find a number δ such that $|f(z) - f(z_0)| < \in$ for all points z of the domain satisfying $|z - z_0| < \delta$. Then $f(z)$ is said to be continuous at $z = z_0$ if $\lim_{z \to z_0} f(z) = f(z_0)$.

Differentiability of a funtion of complex variable

Let $f(z)$ be a single-valued function of the variable z, then

$$f'(z) = \lim_{\delta z \to 0} \frac{f(z + \delta z) - f(z)}{\delta z}$$

provided that the limit exists and is independent of the path along which $\delta z \to 0$. If P be a fixed point and Q be any neighbouring point on the complex plane $z = x + iy$. Then the point Q may approach the point P along any straight line or curved path.

Inverse function corresponding to a funtion of complex variable

If $w = f(z)$, then we may consider z as a function of w, written $z = g(w) = f^{-1}(w)$, where the function f^{-1} is often called the in verse function corresponding to f. Thus, $w = f(z)$ and $w = f^{-1}(z)$ are inverse functions of each other.

Example on finding inverse of a complex function

1. If $f(x) = 2x - 5$, find the inverse.

   Solution :

   Step 1: Change $f(x)$ to $y = 2x - 5$

   Step 2: Switch x and y: $x = 2y - 5$

   Step 3: Solve for y by adding 5 on each side : $\frac{x+5}{2} = y$

   Step 4: Change y back to $f^{-1}(x)$: $f^{-1}(x) = \frac{x+5}{2}$

2. $f(x) = 5x - 7$ and $f^{-1}(x) = \frac{x+7}{5}$.

3. Prove that $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$

   Solution :

   Let $x = \cos\theta + i\sin\theta$, then, $\frac{1}{x} = \cos\theta - i\sin\theta$. Now, $x + \frac{1}{x} = \cos\theta + i\sin\theta +$ $\cos\theta - i\sin\theta = 2\cos\theta$. Taking cube on both sides, $(2\cos\theta)^3 = \left(x + \frac{1}{x}\right)^3 =$ $x^3 + 3\left(x + \frac{1}{x}\right) + x^{-3}$, or, $8\cos^3\theta = \cos 3\theta + i\sin 3\theta + 6\cos\theta + \cos 3\theta - i\sin 3\theta$, or, $8\cos^3\theta = 2\cos 3\theta + 6\cos\theta$, or, $4\cos^3\theta = \cos 3\theta + 3\cos\theta$, or, $\cos 3\theta =$ $4\cos^3\theta - 3\cos\theta$.

4. Prove that sin $\sin 3\theta = 3\sin\theta - 4\sin^3\theta$.

   Solution :

   Let $x = \cos\theta + i\sin\theta$, then, $\frac{1}{x} = \cos\theta - i\sin\theta$. Now, $x - \frac{1}{x} = \cos\theta + i\sin\theta -$

$\cos\theta + i\sin\theta = 2i\sin\theta$. Taking cube on both sides, $(2i\sin\theta)^3 = \left(x - \frac{1}{x}\right)^3 =$

$x^3 - 3\left(x - \frac{1}{x}\right) - x^{-3}$, or, $-8i\sin^3\theta = \cos 3\theta + i\sin 3\theta - 6i\sin\theta - \cos 3\theta + i\sin 3\theta$,

or, $-8i\sin^3\theta = 2i\sin 3\theta - 6i\sin\theta$, or, $-4\sin^3\theta = \sin 3\theta - 3\sin\theta$, or, $\sin 3\theta =$ $3\sin\theta - 4\sin^3\theta$.

5. Prove that $\sinh 3x = 3\sinh x + 4\sinh^3 x$.

   Solution :

   $\sin 3x = 3\sin x - 4\sin^3 x$ or, $\sin 3ix = 3\sin ix - 4\sin^3 ix$ or, $i\sinh 3x = i3\sinh x -$ $4(i\sinh x)^3$ [Since, $\sin iz = i\sinh z$] or, $i\sinh 3x = i3\sinh x + 4i\sinh^3 x$ or, $\sinh 3x = 3\sinh x + 4\sinh^3 x$.

6. Prove that $\cosh 3x = 4\cosh^3 x - 3\cosh x$

   Solution :

   $\cos 3x = 4\cos^3 x - 3\cos x$ or, $\cos 3ix = 4\cos^3 ix - 3\cos ix$ or, $\cosh 3x = 4$ $(\cosh x)^3 - 3\cosh x$ [Since, $\cos iz = \cosh z$] or, $\cosh 3x = 4\cosh^3 x - 3\cosh x$

1.8 Analyticity and Cauchy-Riemann conditions

Objectives

After studying this section, students will be able to-

- understand the basic concepts of analyticity
- learn about Cauchy-Riemann relations and harmonic functions
- find derivatives of complex functions
- learn to identify harmonic functions
- learn applications of the analyticity and Cauchy-Riemann conditions

Introduction

This section aims to define and discuss some of the important properties of complex functions. A function $f(z)$ is aid to be **analytic** if it has a complex derivative $f'(z)$. If the derivative $f'(z)$ exists at all points z of a region R, then $f(z)$ is said to be analytic in R and is referred to as an analytic function in R or a function analytic in R. The terms regular and holomorphic are sometimes used as synonyms for analytic. Thus the function $f(z)$ is said to be analytic at some point z_0 if there exists a **neighborhood** $|z - z_0| < \delta$, at all points of **which** $f'(z)$ exists. In complex analysis we defined the derivative as a limit:

$$\left.\begin{aligned} f'(z) &= \frac{df}{dz} \\ &= \lim_{\Delta z \to 0} \frac{\Delta f}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} \end{aligned}\right\} \qquad (1.74)$$

Analyticity and Cauchy-Riemann conditions

A function $f(z)$ is said to be analytic (holomorphic) at a point $z = z0$ if it is differentiable not only at $z = z_0$ but at every point in a neighborhood of z_0. A function $f(z)$ is analytic in a domain (D) if it is analytic at each point of the domain (D).

A function analytic at every point of C is said to be entire.

If $f(z)$ is analytic in an open domain D, then each of its derivatives $f'(z)$, $f''(z)$, ...exists and is analytic in D.

Derivatives of a complex function

The definition of the complex derivative of a complex function is similar to that of a real derivative of a real function: For a function $f(z)$ the derivative f at z_0 is defined as

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z_0} \qquad (1.75)$$

If the limit exists, we say f is analytic at z_0 or f is differentiable at z_0.

Cauchy-Riemann equations (or conditions)

A necessary condition that $w = f(z) = u(x, y) + iv(x, y)$ be analytic in a region R is that, in R, u and v satisfy the Cauchy - Riemann equations

$$\left.\begin{aligned}\frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x}\end{aligned}\right\} \quad (1.76)$$

If the partial derivatives in Eq. (1.76) are continuous in R, then the Cauchy-Riemann equations are sufficient conditions that $f(z)$ be analytic in R.

The functions $u(x, y)$ and $v(x, y)$ are sometimes called conjugate functions. For a given u which have continuous first order partial derivative on a simply connected region R, we can determine v (within an arbitrary additive constant) so that $u + iv = f(z)$ is analytic.

Derivation of Cauchy–Riemann equations

I. *Necessary condition:*

In order for $f(z)$ to be analytic R, the limit

$$\left.\begin{aligned}&\lim_{\Delta\to 0}\frac{f(z+\Delta z-f(z)}{\Delta z}=f'(z)\\ &=\lim_{\substack{\Delta x\to 0\\ \Delta x\to 0}}\frac{\left[u(x+\Delta x,y+\Delta y)+iv(x+\Delta x,y+\Delta y)\right]-\left[u(x,y)+iv(x,y)\right]}{\Delta x+i\Delta y}\end{aligned}\right\} \quad (1.77)$$

must exist independent of the manner in which Δz (or Δx and Δy) approaches zero. We consider two possible approaches:

Case 1. $\Delta y = 0, \Delta x \to 0$. In this case, Eq. (1.69) becomes

$$\lim_{\Delta x\to 0}\left\{\frac{u(x+\Delta x,y)-u(x,y)}{\Delta x}+i\frac{v(x+\Delta x,y)-v(x,y)}{\Delta x}\right\}=\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}$$

Case 2. $\Delta x = 0, \Delta y \to 0$. In this case, Eq. (1.69) becomes

$$\lim_{\Delta y\to 0}\left\{\frac{u(x,y+\Delta y)-u(x,y)}{i\Delta x}+i\frac{v(x,y+\Delta y)-v(x,y)}{i\Delta x}\right\}=i\frac{\partial u}{\partial y}+\frac{\partial v}{\partial y}$$

subject to the requirement that the partial derivatives exist. Now $f(z)$ cannot possibly be analytic unless these two limits re identical. Thus, a necessary condition that $f(z)$ be analytic is

$$\frac{\partial u}{\partial x}+i\frac{\partial v}{\partial x}=-i\frac{\partial u}{\partial y}+\frac{\partial v}{\partial y}$$

From which we obtain the desired Cauchy-Reimann conditions:

$$\left.\begin{aligned}\frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y}\\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x}\end{aligned}\right\} \tag{1.78}$$

II. *Sufficient condition:*

Since $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ are supposed to be continuous, we have

$$\begin{aligned}\Delta u &= u(x+\Delta x, y+\Delta y)-u(x,y)\\ &= \{u(x+\Delta x, y+\Delta y)-u(x, y+\Delta y)\}+\{u(x, y+\Delta y)-u(x,y)\}\\ &= \left(\frac{\partial u}{\partial x}+\epsilon_1\right)\Delta x+\left(\frac{\partial u}{\partial y}+\eta_1\right)\Delta y\\ &= \frac{\partial u}{\partial x}\Delta x+\frac{\partial u}{\partial y}\Delta y+\epsilon_1\Delta x+\eta_1\Delta y\end{aligned}$$

where $\epsilon_1 \to 0$ and $\eta_1 \to 0$ as $\Delta x \to 0$ and $\Delta y \to 0$

Similarly, since $\frac{\partial v}{\partial x}$ and $\frac{\partial v}{\partial y}$ are supposed to be continuous, we have

$$\begin{aligned}\Delta v &= v(x+\Delta x, y+\Delta y)-v(x,y)\\ &= \{v(x+\Delta x, y+\Delta y)-v(x, y+\Delta y)\}+\{v(x, y+\Delta y)-v(x,y)\}\\ &= \left(\frac{\partial v}{\partial x}+\epsilon_2\right)\Delta x+\left(\frac{\partial v}{\partial y}+\eta_2\right)\Delta y\\ &= \frac{\partial v}{\partial x}\Delta x+\frac{\partial v}{\partial y}\Delta y+\epsilon_2\Delta x+\eta_2\Delta y\end{aligned}$$

where $\epsilon_2 \to 0$ and $\eta_2 \to 0$ as $\Delta x \to 0$ and $\Delta y \to 0$.

Then,

$$\Delta w = \Delta u + i\Delta v$$
$$= \left(\frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \epsilon_1 \Delta x + \eta_1 \Delta y\right)$$
$$+ i\left(\frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y + \epsilon_2 \Delta x + \eta_2 \Delta y\right)$$
$$= \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)\Delta x + \left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right)\Delta y + \epsilon \Delta x + \eta \Delta y$$

where $\epsilon = (\epsilon_1 + i\epsilon_2) \to 0$ and $\eta = (\eta_1 + i\eta_2) \to 0$ as $\Delta x \to 0$ and $\Delta y \to 0$.

By Cauchy-Riemann equations we have

$$\Delta w = \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)\Delta x + \left(-\frac{\partial v}{\partial x} + i\frac{\partial u}{\partial x}\right)\Delta y + \epsilon \Delta x + \eta \Delta y$$
$$= \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)(\Delta x + i\Delta y) + \epsilon \Delta x + \eta \Delta y$$

Then by dividing Δw by $\Delta z = \Delta x + i\Delta y$ and taking the limit $\Delta z \to 0$ (i.e., $\Delta x \to 0, \Delta y \to 0$), we have

$$\lim_{\Delta z \to 0}\frac{\Delta w}{\Delta z} = \frac{df}{dz}$$
$$= f'(z)$$
$$= \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)$$

so that the derivative exists and is unique, i.e., $f(z)$ is analytic in R.

Harmonic functions

If the second partial derivatives of u and v with respect to x and y exist and are continuous in a region R, then from Eq. (1.68) we may find

$$\left.\begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \\ \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0. \end{aligned}\right\} \tag{1.79}$$

It follows that under these conditions, the real and imaginary parts of an analytic function satisfy Laplace's equation denoted by

$$\left.\begin{aligned} \frac{\partial^2\Phi}{\partial x^2}+\frac{\partial^2\Phi}{\partial y^2} &= 0 \\ \text{or, } \nabla^2\Phi &= 0 \\ \text{where, } \nabla^2 &= \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2} \end{aligned}\right\} \tag{1.80}$$

The operator ∇^2 is often called the Laplacian. Functions such as $u(x,y)$ and $v(x,y)$ which satisfy Laplace's equation in a region R are called harmonic functions and are said to be harmonic in R.

Solved problems on analyticity functions and Cauchy-Riemann conditions

1. Find the derivative of $f(z)=z^2$.

Solution :

We may solve this using the definition of the derivative as a limit.

$$f'(z)=\lim_{\Delta z\to 0}\frac{f(z+\Delta z)-f(z)}{\Delta z}=\lim_{\Delta z\to 0}\frac{(z+\Delta z)^2-z^2}{\Delta z}$$

$$=\lim_{\Delta z\to 0}\frac{z^2+2z\Delta z+(\Delta z)^2-z^2}{\Delta z}=\lim_{\Delta z\to 0}(2z+\Delta z)=2z$$

2. Using the definition, find the derivative of $w=f(z)=z^3-2z$ at the point where (a) $z = z_0$, (b) $z = -1$.

Solution :

(a) By definition, the derivative at $z = z_0$ is $f'(z_0)=\lim_{\Delta z\to 0}\frac{f(z_0+\Delta z)-f(z_0)}{\Delta z}$

$$=\lim_{\Delta z\to 0}\frac{\left[(z_0+\Delta z)^3-2(z_0+\Delta z)-\left(z_0^3-2z_0\right)\right]}{}$$

$= \lim_{\Delta z \to 0} \frac{z_0^3 + 3z_0^2\Delta z + 3z_0(\Delta z)^2 + z_0^3 - 2z_0 - 2\Delta z - z_0^3 + 2z_0}{\Delta z} = \lim_{\Delta z \to 0}$

$\frac{3z_0^2\Delta z + 3z_0(\Delta z)^2 - 2\Delta z}{\Delta z} = 3z_0^2 - 2$

In general, $f'(z) = 3z^2 - 2$ for all z.

(b) From (a), or directly, we find that if $z_0 = -1$, then $f'(-1) = 3(-1)^2 - 2 = 1$.

3. Assuming $f(z) = \bar{z}$, show that the limit for $f'(0)$ does not converge.

Solution :

$f'(0) = \lim_{\Delta z \to 0} \frac{f(0+\Delta z) - f(0)}{\Delta z} = \lim_{\Delta z \to 0} \frac{\bar{\Delta z}}{\Delta z} = \lim_{\Delta z \to 0} \frac{\Delta x - i\Delta y}{\Delta x + i\Delta y}$ As, $\Delta z = \Delta x + iy$,

we have $\lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x} = 1$ when $\Delta y = 0$ and $\lim_{\Delta y \to 0} \frac{-i\Delta x}{i\Delta y} = -1$ when $\Delta x = 0$. Since, $\Delta z \to 0$ means both Δx and Δy have to go to 0.

4. (a) Prove that $u(x, y) = e^{-x}(x \sin y - y \cos y)$ is harmonic. (b) Find $v(x, y)$ such that $f(z) = u + iv$ is analytic.

Solution :

(a) $\frac{\partial u}{\partial x} = (-e^{-x})(x \sin y - y \cos y) + (e^{-x})(\sin y)$

$$= -xe^{-x} \sin y + ye^{-x} \cos y + e^{-x} \sin y \tag{1.81}$$

$\frac{\partial^2 u}{\partial x^2} = xe^{-x} \sin y - e^{-x} \sin y - ye^{-x} \cos y - e^{-x} \sin y$

$$= xe^{-x} \sin y - 2e^{-x} \sin y - ye^{-x} \cos y \tag{1.82}$$

$\frac{\partial u}{\partial y} = (e^{-x})(x \cos y + y \sin y - \cos y)$

$$= xe^{-x} \cos y + ye^{-x} \sin y - e^{-x} \cos y \tag{1.83}$$

$\frac{\partial^2 u}{\partial y^2} = (e^{-x})(-x \sin y + y \cos y + \sin y + \sin y)$

$$= -xe^{-x} \sin y + ye^{-x} \cos y + 2e^{-x} \sin y \tag{1.84}$$

Adding (1.72) and (1.74) we get $\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0$. As, u the function $u(x,y)$ satisfies Laplace's equation, it is harmonic.

Solution:

(b) Since $f(z)$ is analytic, we may apply Cauchy-Riemann conditions:

$$\frac{\partial v}{\partial y}=\frac{\partial u}{\partial x}$$

$$=(-e^{-x})(x\sin y-y\cos y)+(e^{-x})(\sin y)$$

$$=-xe^{-x}\sin y+ye^{-x}\cos y+e^{-x}\sin y \tag{1.85}$$

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

$$=(e^{-x})(x\cos y+y\sin y-\cos y)$$

$$=xe^{-x}\cos y+ye^{-x}\sin y-e^{-x}\cos y \tag{1.86}$$

Integrating (1.75), we have

$$v=\int\left(\frac{\partial v}{\partial y}\right)dy$$

$$=\int(-xe^{-x}\sin y+ye^{-x}\cos y+e^{-x}\sin y)dy$$

$$=xe^{-x}\cos y+ye^{-x}\sin y+e^{-x}\cos y-e^{-x}\cos y+F(x)$$

$$=xe^{-x}\cos y+ye^{-x}\sin y+F(x) \tag{1.87}$$

where $F(x)$ is an arbitrary real function of x. Using (1.77) into (1.76) we may obtain

$$-\frac{\partial v}{\partial x}=-\frac{\partial}{\partial x}\left(xe^{-x}\cos y+ye^{-x}\sin y+F(x)\right)$$

$$=xe^{-x}\cos y+ye^{-x}\sin y-e^{-x}\cos y$$

$$\Rightarrow xe^{-x}\cos y-e^{-x}\cos y+ye^{-x}\sin y-F'(x)$$

$$=xe^{-x}\cos y+ye^{-x}\sin y-e^{-x}\cos y \tag{1.88}$$

$$\Rightarrow F'(x)=0$$

$$\Rightarrow F(x)=\text{constant}=c(say) \tag{1.89}$$

Hence $v(x,y)=xe^{-x}\cos y+ye^{-x}\sin y+c$.

5. Show that the function $f(z) = z^2$ is analytic.

Solution :

Let, $f(z) = u(x, y) + iv(x, y)$ Now, $f(z) = z^2 = (x + iy)^2 = (x^2 - y^2) + i(2xy)$ Thus we have, $u(x, y) = x^2 - y^2$ and $v(x, y) = 2xy$

Computing the partial derivatives we have :

$$\frac{\partial u}{\partial x} = 2x, \frac{\partial u}{\partial y} = -2y, \frac{\partial v}{\partial x} = 2y, \frac{\partial v}{\partial y} = 2x$$

Since, the given function satisfies Cauchy-Riemann conditions:

$$\left.\begin{array}{ccccc} \frac{\partial u}{\partial x} & = & \frac{\partial v}{\partial y} & = & 2x \\ \frac{\partial u}{\partial y} & = & -\frac{\partial v}{\partial x} & = & -2y \end{array}\right\} \quad (1.90)$$

we may conclude that $f(z) = z^2$ is analytic.

Alter:

$$f(z) = z^2 \Rightarrow u(x, y) = x^2 - y^2; v(x, y) = 2xy.$$

Applying Cauchy-Riemann sufficient condition we get

$f'(z) = \lim\limits_{\Delta x \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} = \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right) = (2x + i2y) = 2z$. As the derivative exists, the given function is analytic.

6. Check the analyticity of the function defined as

$$f(z) = \begin{cases} \frac{\bar{z}^2}{z}, when\, z \neq 0, \\ 0, when\, z = 0 \end{cases}$$

7. Which of the following is an analytic function of z everywhere in the complex plane?

(a) z^2 (b) $(z^*)^2$ (c) $|z|^2$ (d) $\sqrt{z}$

Ans.: (a)

Solution:

$z^2 = (x+iy)^2 = x^2 - y^2 + i2xy \Rightarrow u(x,y) = x^2 - y^2,\ v(x,y) = 2xy,$ Cauchy Riemann equations : $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = 2x,\ \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} = 2y$ satisfied.

8. What is the value of the contour integral $\oint_C \frac{1}{1+z^2} dz$ evaluated along a contour going from $-\infty$ to ∞ along the real axis and closed in the lower half-plane circle (up to two decimal places)?

Ans: π

Solution :

$$\oint_C \frac{1}{1+z^2} dz = \oint_{-\infty}^{\infty} \frac{1}{1+x^2} dx + \oint_C \frac{1}{1+z^2} dz$$

Poles, $1+z^2 = 0 \Rightarrow z = \pm i, z = -i$ is inside C

$$\text{Res } (z=-i) = \lim_{z \to -i} (z+i) \frac{1}{(z-i)(z+i)} = \lim_{z \to -i} \frac{1}{z-i} = \frac{1}{-i-i} = -\frac{1}{2i}$$

$$\oint_{-\infty}^{\infty} \frac{1}{1+x^2} dx = (-2\pi i) \times \left(-\frac{1}{2i}\right) = \pi$$

(Since here we use lower half-plane i.e., we traversed in the clockwise direction, hence we have to take $-2\pi i$).

1.9 Singular functions

Objectives

After studying this section, students will be able to-

- understand the basic concepts of singularity and their natures
- understand the concepts of poles and branch points
- understand about different types of singularities
- locate singularities of a given function
- find the order of singularity of a given function
- understand the concept of branch cuts

Introduction

In general, a singularity is a point at which an equation, surface, etc., blows up or becomes degenerate. Singularities are often also called singular points. Singularities are extremely important in complex analysis, where they characterize the possible behaviors of analytic functions. Complex singularities are points z_0 in the domain of a function f where f fails to be analytic. Isolated singularities may be classified aspoles, essential singularities, logarithmic singularities, removable singularities, etc. Non-isolated singularities may arise as natural boundaries or branch cuts.

Definition of singular function

If the function $f(z)$ fails to be analytic at some point in the region R, it is then called a singular function. For example the function $f(z)=\frac{1}{z-a}$ fails to be analytic at $z = a$.

A point at which $f(z)$ fails to be analytic is called a singular point or singularity of $f(z)$. For example, the function $f(z)=\frac{1}{z-2}$ has a singular point at the location where $z - 2 = 0$ or $z = 2$.

Types of singularities

There are several types of singularities including- Isolated and nonisolated singularities, Removable singularities, Essential singularities, etc.

Isolated and non-isolated singularities

The point $z = z_0$ is called an isolated singularity or isolated singular point of $f(z)$ if we can find $\delta > 0$ such that the circle $|z - z_0| = \delta$ encloses no singular point other than z_0 (i.e., there exists a deleted neighborhood of z_0 containing no singularity). If no such $\delta > 0$ can be found, we call z_0 a non-isolated singularity.

If z_0 is not a singular point and we can find $\delta > 0$ such that $|z - z_0| = \delta$, encloses no singular point, then we call z_0 an ordinary point of $f(z)$. Removable singularities

An isolated singular point z_0 is called a removable singularity of $f(z)$ if $\lim_{z\to z_0} f(z)$ exists. By defining $f(z_0)=\lim_{z\to z_0} f(z)$, it can then be shown that $f(z)$ is not only continuous at z_0 but is also analytic at z_0.

Example of removable singularity

1. The singular point $z = 0$ is a removable singularity of $f(z) = \frac{\sin z}{z}$ since $\lim_{z \to z_0} \frac{\sin z}{z} = 1$.

Essential singularities

An isolated singularity that is not a pole or removable singularity is called an essential singularity.

Example of essential singularity

1. $f(z) = e^{\frac{1}{z-2}}$ has an essential singularity at $z = 2$.

   If a function has an isolated singularity, then the singularity is either removable, a pole, or an essential singularity. For this reason, a pole is sometimes called a non-essential singularity. Equivalently, $z = z_0$ is an essential singularity if we cannot find any positive integer n such that $\lim_{z \to z_0} (z - z_0)^n f(z) = A \neq 0$.

Singularities at infinity

The type of singularity of $f(z)$ at $z = \infty$ (the point at infinity) is the same as that of $f(1/\omega)$ at $\omega = 0$.

Example of singular function

1. The function $f(z) = z^3$ has a pole of order 3 at $z = \infty$, since $f(1/\omega) = \frac{1}{\omega^3}$ has a pole of order 3 at $\omega = 0$.

Poles and branch points

If z_0 is an isolated singularity and we can find a positive integer n such that $\lim_{z \to z_0} (z - z_0)^n f(z) = A \neq 0$, then $z = z_0$ is called a pole of order n. If $n = 1$, z_0 is called a simple pole.

Examples of poles

1. $f(z)=\dfrac{1}{(z-2)^3}$ has a pole of order 3 at z = 2.

2. $f(z)=\dfrac{3z-2}{(z-1)^2(z+1)(z-4)}$ has a pole of order 2 at z = 1, and simple poles at z = –1 and z = 4.

   If $g(z)=(z-z_0)^n f(z)$, where $f(z_0)$ = 0 and n is a positive integer, then $z = z_0$ is called a zero of order n of $g(z)$. If n = 1, z_0 is called a simple zero. In such a case, z_0 is a pole of order n of the function $\dfrac{1}{g(z)}$.

3. Which of the following statements is TRUE for the function $f(z)=\dfrac{z\sin z}{(z-\pi)^2}$?

   (a) $f(z)$ is analytic everywhere in the complex plane.

   (b) $f(z)$ has a zero at $z = \pi$.

   (c) $f(z)$ has a pole of order 2 at $z = \pi$.

   (d) $f(z)$ has a simple pole at $z = \pi$.

Answer: (c). Reason: $f(z)$ has a pole of order 2 at $z = \pi$.

Branch points

Branch points of multiple-valued functions, are non-isolated singular points since a multiple-valued function is not continuous and, therefore, not analytic in a deleted neighborhood of a branch point.

Branch points are generally the result of a multi-valued function, such as $\sqrt{z}$ or $\log(z)$ being defined within a certain limited domain so that the function can be made single-valued within the domain.

Definition: The point z_0 is called a branch point for the complex (multiple) valued function $f(z)$- if the value of $f(z)$ does not return to its initial value as a closed curve around the point is traced (starting from some arbitrary point on the curve), in such a way that f varies continuously as the path is traced.

Example of branch points

1. $f(z)=\sqrt{z-3}$ has a branch point at $z = 3$.

2. $f(z)=\ln(z^2+z-2)$ has branch points where $\left(z^2+z-2=0\right)$ i.e., at $z = 1$ and $z = -2$.

Order of singularity

The order of singularity is defined in terms of the order of pole of a given function. And a function $f(z)$ can be said to have a pole of order m at z_0 if m is the largest positive integer such that $a_{-m} \neq 0$. A pole of order one is a simple pole. A pole of order two is a double pole, etc.

Example on order of singularity

1. The function $f(z)=\dfrac{1}{(z-3i)^7}$ is singular function, having singularity at $z = 3i$. This function can also be said to have a pole of order 7 at $z = 3i$. So the order of singularity of this function is 7.

Branch cuts

The branch cut is a line or curve excluded from the domain to introduce a technical separation between discontinuous values of the function. When the cut is genuinely required, the function will have distinctly different values on each side of the branch cut.

The shape of the branch cut is a matter of choice, however, it must connect to two different branch points (like $z = 0$ and $z = \infty$ for $\log(z)$) which are fixed in place.

Let us consider the complex valued function

$$\log(z)=\ln(r)+i\theta \tag{1.91}$$

where $z = re^{i\theta}$,, with $r > 0$ and θ real. As one goes around the closed path in Figure 1.3, starting counter-clockwise from point A and returning to A, it is clear that θ increases to $i\theta+2\pi$. Therefore, upon tracing the path, we have:

$$\log(A) = \log(A) + 2\pi i \tag{1.92}$$

This means that $\log(z)$ does not return to its original value when one tries to define it continuously along the closed path. Thus we have an identity crisis: which value should we choose for log(z) at A? Of course, A is arbitrary, so this problem arises at every point in the complex plane!

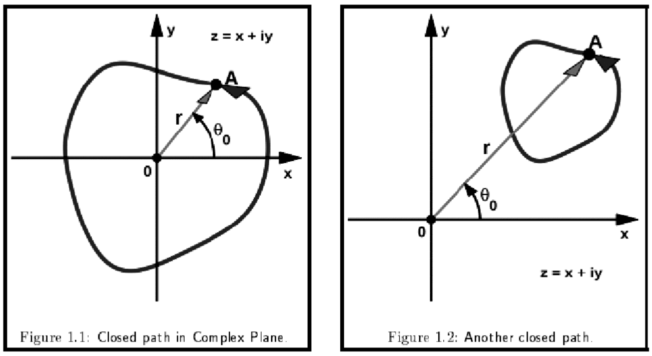


Figure 1.4: Branch-cut

Before answering this question let us first note that θ returns to its original value as z traces the closed path in Figure 1.3. Thus one may ask the question: *what is the difference between the paths in these two figures which makes the behavior of* log(z) *so entirely different as the closed paths are traced?* The answer is that the first path encloses the origin z = 0, while the second path does not. That is why θ increases by 2À as one goes around the second path. Thus the origin is a branch point of log(z).

Cusps

In algebraic geometry, a singularity of an algebraic variety is a point of the variety where the tangent space may not be regularly defined. The simplest example of singularities are curves that cross themselves. But there are other types of singularities, like **cusps**. For example, the equation $y^2 - x^3 = 0$ defines a curve that has a cusp at the origin $x = y = 0$. One could define the x-axis as a tangent at this point, but this definition can not be the same as the definition at other points. In fact, in this case, the x-axis is a “double tangent”.

1.10 Integration of the function of a complex variable

Objectives

After studying this section, students will be able to-

- understand the basic concepts of the complex line integral
- understand the connection between real and complex integrals

- to know the properties of complex integrals
- evaluate complex integrals

Introduction

The concept of the definite integral of real functions does not directly extend to the case of complex functions since real functions are usually integrated over intervals and complex functions are integrated over curves. Interestingly complex integrations are not so complex to evaluate, most of the time easier than the evaluation of real integrations. Some real integrals which are otherwise difficult to evaluate can be evaluated easily by complex integration, and some basic properties of analytic functions are established by complex integration only. The concept of definite integral $\int_a^b f(y)dy$, as studied in the calculus of real-valued function f on a real variable y was generalized to the line integral in vector analysis. Here we extend the concept once more and consider the line integral of a complex function. As in calculus of a real variable, here also we distinguish between definite integrals and indefinite integrals. Complex definite integrals are called the line integrals and are written as $\int_C f(z)dz$. The integrand $f(z)$ is integrated over a given curve C in the complex plane called the path of integration normally represented by a parametric representation $z(t) = x(t) + iy(t), a \le t \le b$. The sense of increasing t is called the positive sense on C which is assumed to be a smooth curve, having continuous derivatives at all $t \in (a,b)$. If the initial and final points of a curve coincide i.e., when $z(a) = z(b)$, the curve is said to be a closed one.

Complex line integrals

Let $f(z)$ be continuous at all points of a curve C (see Fig. 1.5), which we shall assume has a finite length, i.e., C is a rectifiable curve. Subdivide C into n parts by means of points $z_1, z_2, \ldots, z_{n-1}$, chosen arbitrarily, and call $a = z_0$, $b = z_n$. On each arc joining z_{k-1} to z_k [where k goes from 1 to n], choose a point ξ_k. Form the sum

$$S_n = f(\xi_1)(z_1 - a) + f(\xi_2)(z_2 - z_1) + \ldots\ldots + f(\xi_n)(b - z_{n-1}) \tag{1.93}$$

Using the notation $\Delta z_k = z_k - z_{k-1}$, we can write

$$S_n = \sum_{k=1}^{n} f(\xi_k)(z_k - z_{k-1}) = \sum_{k=1}^{n} f(\xi_k)\Delta z_k \tag{1.94}$$

Let the number of subdivisions n increase in such a way that the largest of the chord lengths $|\Delta z_k|$ approaches zero. Then, since $f(z)$ is continuous, the sum S_n

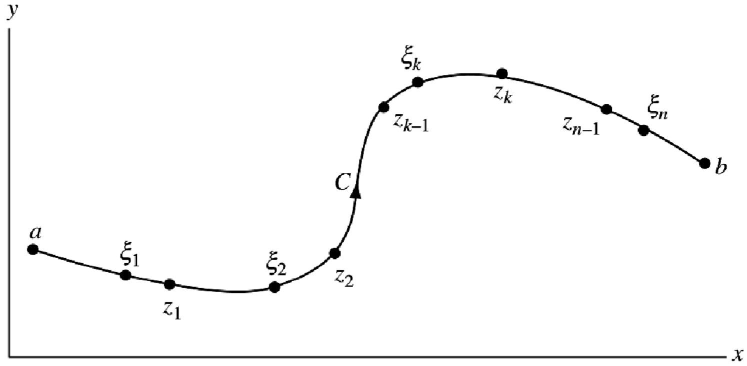


Figure 1.5: Curve on a complex palne.

approaches a limit that does not depend on the mode of subdivision and we denote this limit by

$$\int_a^b f(z)dz \equiv \int_C f(z)dz \tag{1.95}$$

called the complex line integral or simply line integral of $f(z)$ along curve C, or the definite integral of $f(z)$ from a to b along curve C. In such a case, $f(z)$ is said to be integrable along C. If $f(z)$ is analytic at all points of a region R and if C is a curve lying in R, then $f(z)$ is continuous and therefore integrable along C.

Connection between real and complex line integrals

Suppose $f(z) = u(x, y) + iv(x, y) = u + iv$. Then the complex line integral (1.95) can be expressed in terms of real line integrals as follows:

$$\left.\begin{aligned}\int_C f(z)dz &= \int_C (u(x,y)+iv(x,y))(dx+idy)\\ &= \int_C \left[(udx-vdy)+i(vdx+udy)\right]\\ &= \int_C (udx-vdy)+i\int_C (vdx+udy)\end{aligned}\right\} \tag{1.95}$$

Sometimes relation (1.96) is also taken as a definition of a complex line integral.

Properties of integrals

Suppose $f(z)$ and $g(z)$ are integrable along C. Then the following hold:

1. $\int_C (f(z)+g(z))dz = \int_C f(z)dz + \int_C g(z)dz$

2. $\int_C Af(z)dz = A\int_C f(z)dz$, where A = any constant.

3. $\int_a^b f(z)dz = -\int_b^a f(z)dz$

4. $\int_a^b f(z)dz = -\int_a^m f(z)dz + \int_m^b f(z)dz$, where a, b, m are points on C.

5. $\left|\int_C f(z)dz\right| \le ML$, where $|f(z)| \le M$, i.e., M is an upper bound of $|f(z)|$ on C, and L is the length of C.

If P,Q,R are successive points on a curve, property (3) can be written $\int_{PQR} f(z)dz = \int_{RQP} f(z)dz$.

Similarly, if C, C_1, C_2 represent curves from a to b, a to m, and m to b, respectively, it is natural for us to consider $C = C_1 + C_2$ and to write property (4) as $\int_C f(z)dz = \int_{C_1} f(z)dz + \int_{C_2} f(z)dz$.

1.11 Cauchy's integral formula

Objectives

After studying this section, students will be able to-

- learn how the values of an analytic function on a circle determine its values at points enclosed by the circle
- learn applications of the Cauchy integral formula

Introduction

In mathematics, Cauchy's integral formula, named after Augustin-Louis Cauchy, is a central statement in complex analysis. It expresses the fact that a holomorphic function defined on a disk is completely determined by its values on the boundary of the disk, and it provides integral formulas for all derivatives of a holomorphic function. Cauchy's formula shows that, in complex analysis, "differentiation is equivalent to integration": complex differentiation, like integration, behaves well under uniform limits - a result denied in real analysis.

Let $f(z)$ be analytic inside and on a simple closed curve C and let a be any point inside C (see Fig. 1.6). Then

$$f(a) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}dz\,. \qquad (1.97)$$

where C is traversed in the positive (counterclockwise) sense.

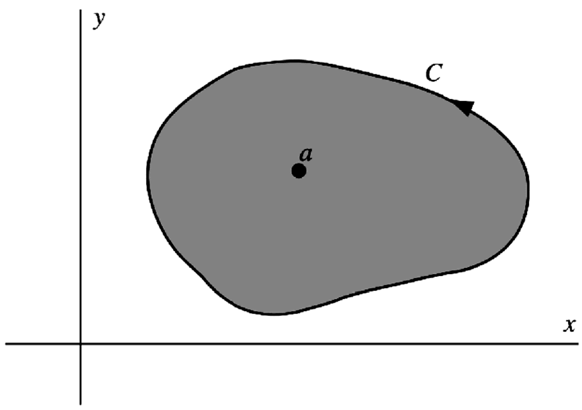


Figure 1.6: Simple closed curve

Proof of Cauchy's Integral formula

Consider the function $\frac{f(z)}{z-a}$, which is analytic at all points within C, except $z = a$. With the point a as centre and radius r, draw a small circle C_1 lying entirely within C.

Now $\frac{f(z)}{z-a}$ is analytic in the region between C and $C1$; hence by Cauchy's integral theorem for multiple connected region we have

$$\begin{aligned}\oint_C \frac{f(z)}{z-a} &= \oint_{C_1} \frac{f(z)}{z-a}dz \\ &= \oint_{C_1} \frac{f(z)-f(a)+f(a)}{z-a}dz \\ &= \oint_{C_1} \frac{f(z)-f(a)}{z-a}dz + f(a)\oint_{C_1} \frac{1}{z-a}dz\end{aligned} \tag{1.98}$$

For any point on C_1

$$\begin{aligned}\oint_{C_1} \frac{f(z)-f(a)}{z-a} &= \int_0^{2\pi}\left[\frac{f(a+re^{i\theta})-f(a)}{re^{i\theta}}\right]ire^{i\theta}d\theta \left[z-a=re^{i\theta}, dz=ire^{i\theta}d\theta\right] \\ &= \int_0^{2\pi}\left[f(a+re^{i\theta})-f(a)\right]id\theta \\ &= 0\left[\text{where } r\to 0\right]\end{aligned}$$

$$\begin{aligned}\oint_{C_1} \frac{1}{z-a}dz &= \int_0^{2\pi}\frac{ire^{i\theta}d\theta}{re^{i\theta}} \\ &= \int_0^{2\pi} id\theta \\ &= i\left[\theta\right]_0^{2\pi} \\ &= 2\pi i\end{aligned}$$

Hence we have,

$$\begin{aligned}\oint_C \frac{f(z)dz}{z-a} &= 0 + f(a)(2\pi i) \\ &= f(a)(2\pi i) \\ \text{or, } f(a) &= \frac{1}{2\pi i}\oint_C \frac{f(z)dz}{z-a}\end{aligned}$$

Also, the nth derivative of $f(z)$ at $z = a$ is given by

$$\left.\begin{aligned}f^{(n)}(a) &= \frac{n!}{2\pi!}\oint_C \frac{f(z)}{(z-a)^{n+1}}dz;\ n=0,1,2,\ldots\ldots \\ \oint_C \frac{f(z)dz}{(z-a)^{n+1}} &= \frac{2\pi!}{n!}f^{(n)}(a)\left[f^{(n)}(a)=\frac{d^n}{dz^n}f(z)\Big|_{z=a}\right]\end{aligned}\right\} \tag{1.99}$$

The result given in Eq.(1.97) can be considered a special case of result in Eq.(1.99) with $n = 0$ if we define $0! = 1$.

The results given in Eq.(1.97) and Eq. (1.99) are called Cauchy's integral formulas and are quite remarkable because they show that if a function $f(z)$ is known on the simple closed curve C, then the values of the function and all its derivatives can be found at all points inside C. Thus, if a function of a complex variable has a first derivative, i.e., is analytic, in a simply-connected region R, all its higher derivatives exist in R. This is not necessarily true for functions of real variables.

Cauchy's Inequality

Suppose $f(z)$ is analytic inside and on a circle C of radius r and centered at $z = a$. Then

$$\left|f^{(n)}(a)\right| \le \frac{M.n!}{r^n}, n = 0,1,2,... \quad (1.100)$$

where M is a constant such that $|f(z)| < M$ on C, i.e., M is an upper bound of $|f(z)|$ on C. We have by Cauchy's integral formulas,

$$f^{(n)}(a) = \frac{n!}{2\pi !}\oint_C \frac{f(z)}{(z-a)^{n+1}}dz, n = 0,1,2,... \quad (1.101)$$

Since, $|z - a| = r$ lies on C and the length of C is $2\pi r$,

$$\left.\begin{aligned} \left|f^{(n)}(a)\right| &= \frac{n!}{2\pi i}\left|\oint_C \frac{f(z)}{(z-a)^{n+1}}dz\right| \\ &\le \frac{n!}{2\pi i}\frac{M}{r^{n+1}}2\pi r \\ &= \frac{M.n!}{r^n} \end{aligned}\right\} \quad (1.102)$$

Solved examples on Cauchy's integral formulas

1. If $f(z)$ be analytic inside and on the boundary C of a simplyconnected region R, then prove Cauchy's integral formula

$$f(a) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}dz. \quad (1.103)$$

Proof: Method 1. The function $f(z) = (z - a)$ is analytic inside and on C except at the point $z = a$.

Theorem 1.11.1 *If $f(z)$ be analytic in a region bounded by two simple closed curves C and C1 (where C1 lies inside C as in Fig. 1.7(a)) and on these curves, then $\oint_C f(z)dz = \oint_{C_1} f(z)dz$ where C and C1 both are traversed in the positive sense (counter- clockwise)* relative to their interiors. The result shows that if we wish

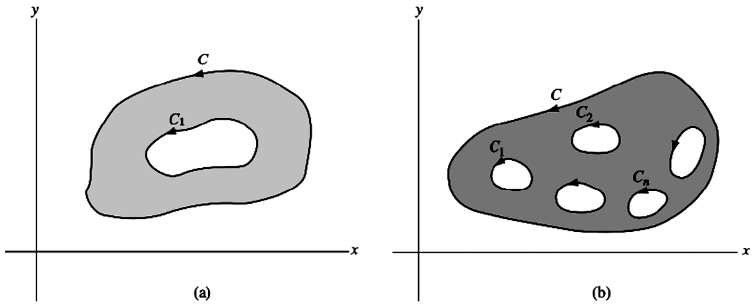
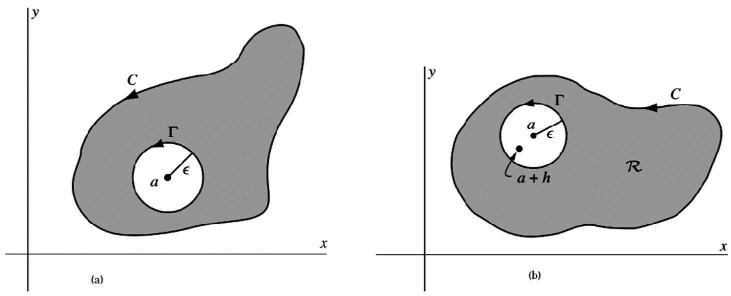


Figure 1.7: Simple closed curves

Figure 1.8: Simply connected regions



to integrate $f(z)$ along curve C, we can equivalently replace C by any curve C_1 so long as $f(z)$ is analytic in the region between C and C_1, as in Fig 1.7(b).

We have

$$\oint_C \frac{f(z)}{z-a}dz = \oint_\Gamma \frac{f(z)}{z-a}dz \qquad (1.104)$$

Where we can choose Γ as a circle of radius $\in$ with center at a. Then an equation for Γ is $|z-a| = \in$ or $z-a = \in e^{i\theta}$ where $0 \le \theta \le \pi$. Substituting $z = a + \in e^{i\theta}$, $dz = \in e^{i\theta} d\theta$ integral on the right of (1.104) becomes

$$\oint_\Gamma \frac{f(z)}{z-a}dz = \int_0^{2\pi} \frac{f(a+\in e^{i\theta})i \in e^{i\theta}}{\in e^{i\theta}} d\theta = i\int_0^{2\pi} f(a+\in e^{i\theta})d\theta$$

From (1.104) we get,

$$\oint_C \frac{f(z)}{z-a}dz = i\int_0^{2\pi} f(a+\in e^{i\theta})d\theta \tag{1.105}$$

Taking the limit of both sides of (1.104) and making use of the continuity of $f(z)$, we have

$$\left.\begin{aligned}
\oint_C \frac{f(z)}{z-a} &= \lim_{\in\to 0'} \int_0^{2\pi} f(a+\in e^{i\theta})d\theta \\
&= i\int_0^{2\pi} \lim_{\in\to 0} \int (a+\in e^{i\theta})d\theta \\
&= i\int_0^{2\pi} \int (a)d\theta \\
&= i(2\pi)f(a) \\
\Rightarrow f(a) &= \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}dz
\end{aligned}\right\} \tag{1.106}$$

2. Evaluate $\frac{1}{2\pi i}\oint_C \frac{f(z)}{z-a}dz$.

Simply and multiply connected region

A region R is called simply-connected if any simple closed curve, which lies in R, can be shrunk to a point without leaving R. A region R, which is not simply-connected, is called multiply- connected. For example, suppose R is the region defined by $|z| < 2$ shown shaded in Fig. 1.9(a). If Γ is any simple closed curve lying in R [i.e., whose points are in R], we see that it can be shrunk to a point that lies in R, and thus does not leave R so that R is simply-connected. On the other hand, if R is the region defined by $1 < |z| < 2$, shown shaded in Fig. 1.9(b), then there is a simple closed curve Γ lying in R that cannot possibly be shrunk to a point without leaving R so that R is multiply-connected. Intuitively, a simply-connected region is

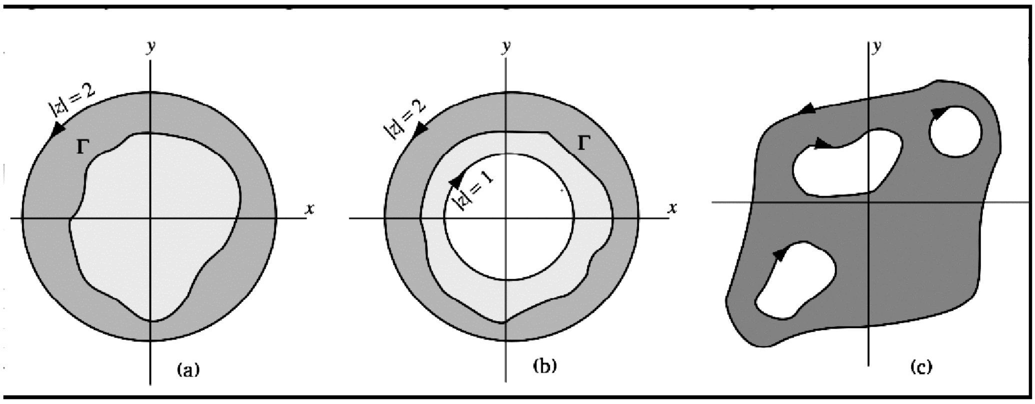


Figure 1.9: Simply and multiply connected regions.

one that does not have any "holes" in it, while a multiply-connected region has "hole(s)". The multiply-connected regions of Fig. 1.9(c) have, respectively, one and three holes in them.

1.12 Laurent and Taylor's expansion

Objectives

After studying this section, students will be able to-

- learn to determine Taylor's series and Laurent's series of some given functions within specified regions
- understand the definition of zeros of order n of an analytic function

Introduction

Taylor's theorem states that any function satisfying certain conditions can be expressed as a Taylor series. And the Laurent series of a complex function f (z) is a representation of that function as a power series which includes terms of negative degree. It may be used to express complex functions in cases where a Taylor series expansion cannot be applied.

Taylor's theorem

Let $f(z)$ be analytic inside and on a simple closed curve C. Let a and $a + h$ be two points inside C. Then

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots + \frac{h^n}{n!}f^{(n)}(a) + \ldots \qquad (1,107)$$

or writing $z = a + h$ or $h = z - a$

$$f(z) = f(a) + f'(a)(z-a) + \frac{f''(a)}{2!}(z-a)^2 + ... + \frac{f^{(n)}(a)}{n!}(z-a)^n + ... \quad (1.108)$$

This is called Taylor's theorem and the series (1.107) or (1.108) is called a Taylor series or expansion for $f(a+h)$ or $f(z)$. The region of convergence of the series (1.108) is given by $|z - a| < R$, where the radius of convergence R is the distance from a to the nearest singularity of the function $f(z)$. On $|z-a| = R$, the series may or may not converge. For $|z - a| > R$, the series diverges. If the nearest singularity of $f(z)$ is at infinity, the radius of convergence is infinite, i.e., the series converges for all z. If $a = 0$ in (1.107) or (1.108), the resulting series is often called a **Maclaurin series**.

Laurent's theorem

Let C_1 and C_2 be concentric circles of radii R_1 and R_2, respect- ively, and center at a (see Fig. 1.10). Suppose that $f(z)$ is single- valued and analytic on C_1 and C_2 and, in the ring-shaped region R [also called the annulus or annular region] between C_1 and C_2, is shown shaded in Fig. 1.10. Let $a + h$ be any point in R. Then we have

$$f(a+h) = a_0 + a_1 h + a_2 h^2 + ... + \frac{a_{-1}}{h} + \frac{a_{-2}}{h^2} + \frac{a_{-3}}{h^3} + ... \quad (1.109)$$

where

$$\left.\begin{aligned} a_n &= \frac{1}{2\pi i}\oint_{C_1}\frac{f(z)}{(z-a)^{n+1}}dz;\ n = 0,1,2,... \\ a_{-n} &= \frac{1}{2\pi i}\oint_{C_1}(z-a)^{n-1}f(z)dz;\ n = 1,2,... \end{aligned}\right\} \quad (1.110)$$

C_1 and C_2 being traversed in the positive direction with respect to their interiors. In the above integrations, we can replace C_1 and C_2 by any

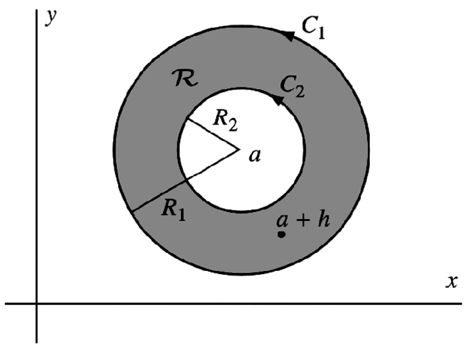


Figure 1.10: Contour

concentric circle C between C_1 and C_2. Then, the coefficients can be written in a single formula,

$$a_n = \frac{1}{2\pi i}\oint_{C_1} \frac{f(z)}{(z-a)^{n+1}} dz; \quad n = 0, \pm 1, \pm 2, \ldots \tag{1.111}$$

With an appropriate change of notation, we can write the above as

$$f(z) = a_0 + a_1(z-a) + a_2(z-a)^2 + \ldots + \frac{a_{-1}}{z-a} + \frac{a_{-2}}{(z-a)^2} + \ldots\ldots \tag{1.112}$$

where

$$a_n = \frac{1}{2\pi i}\oint_{C_1} \frac{f(\varsigma)}{(\varsigma-a)^{n+1}} d\varsigma; \quad n = 0, \pm 1, \pm 2, \ldots \tag{1.113}$$

This is called Laurent's theorem and (1.109) or (1.112) with coefficients (1.110), (1.111), or (1.113) is called a Laurent series or expansion. The part $a_0 + a_1(z - a) + a_2(z - a)^2 + \ldots$ is called the analytic part of the Laurent series, while the remainder of the series, which consists of inverse powers of $(z - a)$, is called the principal part. If the principal part is zero, the Laurent series reduces to a Taylor series.

Problems on Lorentz and Taylor's expansion

1. (a) Expand $f(z) = \sin z$ in a Taylor series about $z = \frac{\pi}{4}$.

   (b) Determine the region of convergence of this series.

2. Expand $f(z) = \frac{1}{(z+1)(z+3)}$ in a Laurent series valid for

   (a) $1 < |z| < 4$

   (b) $|z| > 3$

   (c) $0 < |z + 1| < 2$

   (d) $|z| < 1$

1.13 Residues and Residue Theorem

Objectives

After studying this section, students will be able to-

- understand the concepts of residues
- determine residue at specified poles
- understand the statement of Cauchy's residue theorem
- learn the application of Cauchy's residue theorem

Residue

Let $f(z)$ be single-valued and analytic inside and on a circle C except at the point $z = a$ chosen as the center of C. Then $f(z)$ has a Laurent series about $z = a$ given by

$$\left.\begin{aligned} f(z) &= \sum_{\infty}^{\infty} a_n (z-a)^n \\ &= a_0 + a_1(z-a) + a_2(z-a)^2 + \ldots + \frac{a_{-1}}{z-a} + \frac{a_{-2}}{(z-a)^2} + \ldots\ldots \end{aligned}\right\} \quad (1.114)$$

where

$$a_n = \frac{1}{2\pi i}\oint_C \frac{f(z)}{(z-a)^{n+1}} dz;\ n = 1, \pm 1, \pm 2, \ldots \quad (1.115)$$

In the special case n = "1, we have from (1.115)

$$\oint_C f(z) dz = 2\pi i a_{-1} \quad (1.116)$$

Formally, we can obtain (1.116) from (1.114) by integrating term by term and using the results

$$\oint_C \frac{f(z)}{(z-a)^p} dz = \begin{cases} 2\pi i,\ f \text{ or } p = 1; \\ 0,\ f \text{ or } p = \text{integer} \neq 1 \end{cases} \quad (1.117)$$

Because of the fact that (1.116) involves only the coefficient a_{-1} in (1.114), we call a_{-1} the residue of $f(z)$ at $z = a$.

Calculation of Residues

To obtain the residue of a function $f(z)$ at $z = a$, it may appear from (1.114) that the Laurent expansion of $f(z)$ about $z = a$ must be obtained. However, in the case where $z = a$ is a pole of order k, there is a simple formula for a"1 given by

$$a_{-1} = \lim_{z\to a} \frac{1}{(k-1)!} \frac{d^{k-1}}{dz^{k-1}} \left\{(z-a)^k f(z)\right\} \tag{1.118}$$

If $k = 1$ (simple pole), then the result is especially simple and is given by

$$a_{-1} = \lim_{z\to a} (z-a) f(z) \tag{1.119}$$

which is a special case of (1.118) with $k = 1$ if we define $0! = 1$.

Application of residue in solving definite integrals

1. Find the value of the integral $\oint_C \frac{e^z \sin(z)}{z} dz$, where the contour C is the unit circle: $|z - 2| = 1$.

   Solution:

   $|z-2| = 1 \Rightarrow 1 < z < 2$ i.e. the pole $z = 0$ does not lie inside the contour.

   $$\oint_C \frac{e^z \sin(z)}{z} dz = 2\pi i \times \Sigma(\text{Residue}) = 2\pi i \times 0 = 0$$

   [Residue at the pole $z = 0$ is $\lim_{z\to 0}(z-0) f(z) = \lim_{z\to 0}(z)\left(\frac{e^2 \sin(z)}{z}\right) = \lim_{z\to 0}$ $e^z \sin(z) = 0$]

2. Consider a counterclockwise circular contour $|z| = 1$ about the origin. Let $\frac{z\sin(z)}{z-\pi}$, then the integral $\oint f(z)dz$ over this contour is (a) $-i\pi$ (b) Zero (c) $i\pi$ (d) $2i\pi$

   Solution:

   Since, pole $z = \pi$ does not lie inside the contour, hence

   $$\oint f(z)dz = (2\pi i)\Sigma(\text{Res}) = (2\pi i)\Sigma(0) = 0$$

3. The value of the integral $\oint_C \frac{z^2}{e^z + 1} dz$ where C is the circle $|z| = 4$ is (a) $2\pi i$ (b) $2\pi^2 i$ (c) $4\pi^3 i$ (d) $4\pi^2 i$

Solution:

Pole, $e^z = -1 = e^{(2n+1)i\pi}, n = 0,1,2,3,\ldots$

For $z = i\pi$, $\text{Res} = \lim_{z \to i\pi} \frac{\phi(z)}{\phi'(z)} = -\frac{\pi^2}{e^{i\pi}} = \pi^2$

For $z = -i\pi$, $\text{Res} = \lim_{z \to i\pi} \frac{\phi(z)}{\phi'(z)} = -\frac{\pi^2}{e^{-i\pi}} = \pi^2$

$$\oint_C \frac{z^2}{e^z + 1} dz = 2\pi i \Sigma \,\text{Res} = (2\pi i)(\pi^2 + \pi^2) = 4\pi^3 i$$

Chapter-end exercise

1. Assume: $z_1 = 3i$ and $z_2 = 2 - 2i$.

   (a) Plot the points $z_1 - z_2$, $z_1 + z_2$, and $\bar{z}$.

   (b) Compute $|z_1 + z_2|$ and $|z_1 - z_2|$.

   (c) Express z_1 and z_2 in polar form.

2. Let $z_1 = 6e^{i\pi/3}$ and $z_2 = 2e^{-i\pi/6}$. Plot z_1z_2, and z_1/z_2.

3. Find an identity for sin 3θ using n = 3 in De Moivre's formula. Write your identity in a way that involves only sin θ and sin 3θ if possible.

4. Evaluate $\int_C z^2 dz$ where C is any curve joining 0 and $1 + i$.

5. Find the residue of $f(z) = \frac{z^2 - 1}{z^2 + z}$ at $z = 0$.

6. Using Cauchy's integral formula evaluate $\int_C \frac{z^2}{z-2} dz$ where C is the circle $|z| = 3$.

7. Find an analytic function $f(z) = u + iv$ whose real part is $e^x(s\cos y - y\sin y)$.

8. Find a representation for the function

$$f(z)=\frac{1}{1+z}=\frac{1}{z}.\frac{1}{1+\left(\frac{1}{z}\right)}$$

in negative powers of z that is valid when $1<|z|<\infty$.

9. Find the Taylor series for the function

$$\frac{1}{z}=\frac{1}{2+(z-2)}=\frac{1}{2}.\frac{1}{1+(z-2)/2}$$

about the point $z_0 = 2$. Then by differentiating the series term by term, prove that

$$\frac{1}{z^2}=\frac{1}{4}\sum_{n=0}^{\infty}(-1)^n(n+1)\left(\frac{z-2}{2}\right)^n;\ (|z-2|<2).$$

10. Show that the singular point of each of the following functions is a pole. Determine the order m of that pole and the corresponding residue B.

(a) $f(z)=\frac{1-\exp(2z)}{z^4}$

(b) $f(z)=\frac{1-\cosh z}{z^3}$

(c) $f(z)=\frac{\exp(2z)}{(z-1)^2}$

[Ans. (a) $m=3, B=-\frac{4}{3}$; (b) $m-1, B=-\frac{1}{2}$; (c) $m=2, B=2e^2$]

Unit 2 ❑ Integrals transforms

Structure

2.1 Objectives

2.2 Introduction

2.3 Fourier Transforms

2.4 Representation of Dirac delta function as a Fourier Integral

2.5 Fourier transform of derivatives

2.6 Inverse Fourier transform

2.7 Convolution theorem

2.8 Properties of Fourier transforms

2.9 Three-dimensional Fourier transforms with examples

2.10 Application of Fourier Transforms

2.1 Objective

After successful completion of this chapter, learners will be able to—

- understand the concepts of two highly powerful tools for solving differential equations: Fourier Transforms and Laplace transforms
- understand Fourier integral theorem and derivation of Fourier integral
- understand the basic concepts of Fourier transform
- calculate the Fourier transform of periodic functions including the cosine, sine, and other functions
- learn the Fourier representations of Dirac delta function
- express a convolution mathematically and explain its function
- understands the Fourier transform of a delta function and a shifted delta function.
- learn the Fourier representations of derivatives and differential equations etc
- understand the Fourier inversion theorem

2.2 Introduction

Integral transforms refer to two highly powerful tools for solving dif-ferential ferential equations: i) The Fourier Transforms and ii) The Laplace transforms. Besides practical applications, the Fourier transform has a fundamental importance in quantum mechanics, providing the correspondence between the position and the momentum representations of the Heisenberg commutation relations. An integral transform becomes useful only when it allows one to transform a complicated problem into a much simpler one. hese transforms are also useful for solving integral equations. When one attempts to solve a differential equation, with an unknown function *f*, he/she first applies the transform to the differential equation to turn it into an easily solvable equation: often an algebraic equation for the transform F of *f*. One then solves the resulting equation for F and finally applies the inverse transform to get back *f*. The idea can be sequentially presented as in the block diagram shown in Fig 2.1. It is to be noted that a direct solution from the 1st step to the last step is often difficult. So we have to follow the apparently long but effectively straightforward path. The ultimate purpose of developing such formalism is to reduce the solution of complicated problems to a set of simple rules which even a machine could follow.

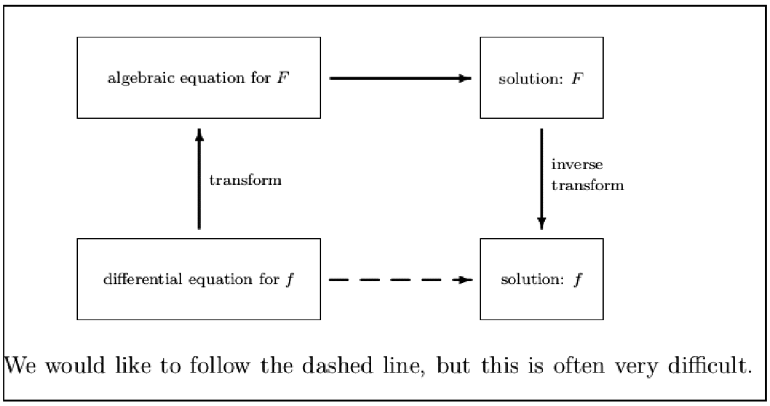


Figure 2.1 : Steps followed in Fourier Transform

2.3 Fourier Transforms

The readers are assumed to have ideas on how to expand a periodic function as a trigonometric series. Which can be thought of as a decomposition of a periodic function in terms of elementary modes, each of which has a definite frequency allowed by the periodicity. If the function has period L, then the frequencies must be integer multiples of the fundamental frequency $k = 2\pi / L$ $\left(i.e., K_n = \frac{2\pi}{L}\right)$. Thus, the Fourier series expansion of a periodic function f(x) with fundamental period L is given by

$$f(x)=\sum_{-\infty}^{\infty}a_n e^{i2\pi nx/L} \tag{2.1}$$

where the coefficients of the series a_n ared by

$$a_n=\frac{1}{L}\int_0^L f(x)e^{-ik_n x}dx \tag{2.2}$$

In this section, we would like to establish a similar decomposition for functions that are not periodic. A non-periodic function can be thought of as a periodic function in the limit $L\to\infty$. The larger L is, the less frequently the function repeats until in the limit $L\to\infty$ the function does not repeat at all. In the limit, $L\to\infty$ the allowed frequencies become a continuum and the Fourier sum goes over to a Fourier integral.

Thus it is understood that Fourier Transform is a tool that breaks a waveform (a function or signal) into an alternate representation, characterized by sine and cosines. And it shows that any waveform can be re-written as the sum of sinusoidal functions.

Fourier Integral theorem

A periodic function $f(x)$ having a fundamental period L, defined on the real line, can be expanded in a Fourier series (converging to almost everywhere within each period) as

$$f(x)=\sum_{-\infty}^{\infty}a_n e^{i2\pi nx/L} \tag{2.3}$$

Where the coefficents a_n are given by

$$a_n=\frac{1}{L}\int_{-l/2}^{L/2} f(x)e^{-i2\pi nx/L}dx. \tag{2.4}$$

Here we have chosen the period to be $[-L/2, L/2]$ for convenience. In the cases where $f(x)$ is not periodic, we can still define a function

$$f_L(x)=\sum_{-\infty}^{\infty}a_n e^{\frac{i2\pi nx}{L}} \tag{2.5}$$

with the same a_n as above. By construction, this function $f_L(x)$ is periodic with period L and moreover agrees with $f(x)$ for almost all $x\in[-L/2, L/2]$. Then it is clear that as we make L larger and larger, then $f_L(x)$ and $f(x)$ agree (almost everywhere) on a larger and larger subset of the real line. One should expect that in the limit $L \to \infty$, $f_L(x)$ should converge to $f(x)$ in some sense. The task ahead is to find reasonable expressions for the limit $L \to \infty$ of the expression (2.5) of $f_L(x)$ and of the coefficients (2.4).

Derivation of Fourier integral

Let $f(x)$ has the following properties:

- understan
- **Prop 1 :** $f(x)$ is piecewise continuous on every interval $[-L/2, L/2]$
- **Prop 2 :** $f(x)$ is absolutely integrable on the x-axis, that is $\int_{-\infty}^{\infty} |f(x)dx|$ converges.

Prof 3 : At every x onthe real line, $f(x)$ has left and right hand derivatives.

Consider the Fourier seies representation of $f(x)$ i on the interval $[-L/2, L/2]$. We have

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n \cos\left(\frac{n\pi x}{L/2}\right) + b_n \sin\left(\frac{n\pi x}{L/2}\right)\right] \tag{2.6}$$

where

$$a_0 = \frac{1}{\frac{L}{2}}\int_{-L/2}^{L/2} f(x')dx' = \frac{2}{L}\int_{-L/2}^{L/2} f(x')dx' \tag{2.7}$$

$$a_n = \frac{1}{L/2}\int_{-L/2}^{L/2} f(x')\sin\left(\frac{n\pi x'}{L/2}\right)dx'$$

$$= \frac{2}{L}\int_{-L/2}^{L/2} f(x')\cos\left(\frac{2n\pi x'}{L}\right)dx' \tag{2.8}$$

$$b_n = \frac{1}{L/2}\int_{-L/2}^{L/2} f(x')\sin\left(\frac{n\pi x'}{L/2}\right)dx'$$

$$= \frac{2}{L}\int_{-L/2}^{L/2} f(x')\sin\left(\frac{2n\pi x'}{L}\right)dx' \tag{2.9}$$

Set $\omega_n = \dfrac{n\pi}{L/2} = \dfrac{2n\pi}{L}$ and

$$\Delta\omega = \omega_n - \omega_{n-1} = \frac{n\pi}{L/2} = \frac{2\pi}{L}(n-n+1) = \frac{2\pi}{L}$$

Then, we can write

$$\left.\begin{aligned} f(x) = &\ \frac{\Delta\omega}{2\pi}\int_{L/2}^{L/2} f(x')dx' \\ &+\frac{\Delta\omega}{\pi}\sum_{n=1}^{\infty}\left[\left\{\int_{-L/2}^{L/2} f(x)\cos(\omega_n x')dx'\right\}\cos(\omega_n x)\right] \\ &+\frac{\Delta\omega}{\pi}\left[\left\{\int_{-L/2}^{L/2} f(x')\sin(\omega_n x')dx'\right\}\sin(\omega_n x)\right] \end{aligned}\right\} \quad (2.10)$$

Let $L \to \infty$, so that $(-L/2, L/2) \to (-\infty, \infty)$. then $\Delta\omega \to 0$, and $\frac{\Delta\omega}{2\pi}\int_{-L/2}^{L/2} f(x')dx' \to 0$, $\int_{-\infty}^{\infty} |f(x)| dx$ converges, Henece, summation can be replaced by integration (as $f(x)$ satisfies the proprties **Propl to prop3** and we have

$$\left.\begin{aligned} f(x) = &\ 0 + \frac{i}{\pi}\int_0^{\infty} d\omega\left[\left\{\int_{-\infty}^{\infty} f(x')\cos(\omega x')dx'\right\}\cos(\omega x)\right] \\ &+\frac{1}{\pi}\int_0^{\infty} d\omega\left[\left\{\int_{-\infty}^{\infty} f(x')\sin(\omega x')dx'\right\}\sin(\omega x)\right] \end{aligned}\right\} \quad (2.11)$$

The above equation in analogous to the Fourier series, wherein the sums are replaced by the integrals. Let us denote :

$$\left.\begin{aligned} A(\omega) &= \int_{-\infty}^{\infty} f(x')\cos(\omega x')dx' \\ B(\omega) &= \int_{-\infty}^{\infty} f(x')\sin(\omega x')dx' \end{aligned}\right\} \quad (2.12)$$

Then from Eq. (2.11) we write

$$f(x) = \frac{1}{\pi}\int_{\infty}^{\infty} d\omega\left[A(\omega)\cos(\omega x) + B(\omega)\sin(\omega x)\right] \quad (2.13)$$

This equation is called the Fourier integral of $f(x)$ or Fourier integral representation of $f(x)$.

Theorem : If f(x) satisfies the properties propl to prop3, then the Fourier integral representation of f(x) conveges to f(x) at a point of continuity and to $\left[\frac{f(x+)+f(x-)}{2}\right]$ at a point of discontinuity.

Fourier integral transform

Fourier transform of the function $f(x)$ is defined as

$$F(k) = F\left[f(x)\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx \tag{2.14}$$

subject to the requirement that the integral exists. Is is to be noted that *not every function $f(x)$ has a Fourier transform.*

A sufficient condition for a function that it has a Fourier transform is that it must be square-integrable or the following integral converges:

$$\|f\|^2 \equiv \int_{-\infty}^{\infty} |f(x)|^2 dx \tag{2.15}$$

Explanation with case study :

Let $f(x) = \frac{1}{4+x^2}$.

Let us evaluate the integral :

$$\left.\begin{aligned} &= \int_{-\infty}^{\infty} |f(x)|^2 dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\left(4+x^2\right)^2} dx \\ &= \frac{\pi}{16} \end{aligned}\right\}$$

[Hints : Let $x = 2\tan\theta$, $dx = 2\sec^2\theta d\theta$, $x[-\infty,\infty] \to \theta\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$]

As the function is square-integrable, its Fourier transform exists and we have :

$$F(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-ikx}}{4+x^2}dx$$

The function $\frac{e^{ikz}}{4+z^2}$ has simple poles at $z = \pm 2i$, According to residue theorem, for $k < 0$, we have to pick up the residues of the poles in the upper half-plane :

$$\left.\begin{aligned} F(k) &= \left(\frac{1}{2\pi}\right)(2\pi i)\sum \text{Res} \\ &= \left(\frac{1}{2\pi}\right)(2\pi i)\text{Res}(2i),\ if\ k \le 0 \\ &= \frac{1}{4}e^{2k} \end{aligned}\right\}$$

and for $k > 0$ we have to pick up the poles in the lower half-plane. Therefore we have

$$\left.\begin{aligned} F(k) &= \left(\frac{1}{2\pi}\right)(2\pi i)\sum \text{Res} \\ &= \left(\frac{1}{2\pi}\right)(2\pi i)\text{Res}(-2i),\ if\ k \le 0 \\ &= \frac{1}{4}e^{2k} \end{aligned}\right\}$$

Therefore,

$$F(k) = \frac{1}{4}e^{-2|k|}\ [k = -\infty, \infty]$$

Fourier sine transform of $f(x)$:

$$F_s[f(x)] = \sqrt{\frac{2}{\pi}}\int_0^{\infty} f(x)\sin(kx)dx \tag{2.16}$$

Finite Fourier sine transform of $f(x)$:

$$F_s[f(x)] = \int_0^{b} f(x)\sin(kx)dx \tag{2.17}$$

$$if,\ f(x) = \begin{cases} 1, 0 < x < b \\ 0,\ x > b \end{cases}$$

$$F[f(x) = \int_0^b f(x)\sin(kx)dx + \int_b^\infty f(x)\sin(kx)dx$$

$$= \int_0^b (1)\sin(kx)dx + \int_b^\infty (0)\sin(kx)dx$$

$$= \frac{1-\cos bk}{k} \qquad (2.18)$$

Examples of Fourier transforms

1. Find the Fourier integral representation of the function

$$f(x) = \begin{cases} 0, if\, x < 0, \\ 1, if\ 0 \le x \le 1, \\ 0\ if\ x > 1. \end{cases}$$

Hence, show that $\int_0^\infty \frac{\sin(x/2)}{x}dx = \frac{\pi}{2}$.

Solution : As the given function satisfies the hypothesis of the theorem, we have

$$A(\omega) = \Rightarrow \int_0^\infty d\omega \left[\frac{1}{x}\sin(x/2)\right] = \frac{\pi}{2}$$

$$= \int_{-\infty}^\infty \cos(\omega x)dx$$

$$= \left[\frac{\sin(\omega x)}{\omega}\right]_0^1$$

$$= \frac{\sin\omega}{\omega}$$

$$B(\omega) = \int_{-\infty}^\infty f(x)\cos(\omega x)dx = \int_{-\infty}^\infty \sin(\omega x)dx$$

$$= \left[\frac{1-\cos(\omega x)}{\omega}\right]_0^1$$

$$= \frac{1-\cos(\omega x)}{\omega}$$

$$f(x) = \frac{1}{\pi}\int_0^\infty d\omega\left[A(\omega)\cos(\omega x)+B(\omega)\sin(\omega x)\right]$$

$$= \frac{1}{\pi}\int_0^\infty d\omega\left[\frac{\sin\omega}{\omega}\cos(\omega x)+\frac{1-\cos\omega}{\omega}\sin(\omega x)\right]$$

$$= \frac{1}{\pi}\int_0^\infty d\omega\left[\frac{2\sin(\omega/2)\cos(\omega/2)\cos(\omega x)+2\sin^2(\omega/2)\sin(\omega x)}{\omega}\right]$$

$$= \frac{2}{\pi}\int_0^\infty d\omega\frac{1}{\omega}\sin(\omega/2)\left[\cos(\omega/2)\cos(\omega x)+\sin(\omega/2)\sin(\omega x)\right]$$

$$\text{or, } f(x) = \frac{2}{\pi}\int_0^\infty d\omega\left[\frac{1}{\omega}\sin(\omega/2)\cos\omega(x-1/2)\right]$$

This is the Fourier representation of the given function.

Let $x = 1/2$, then $f(1/2) = 1$.

$$\text{Hence } 1 = \frac{2}{\pi}\int_0^\infty dw\left[\frac{1}{\omega}\sin(\omega/2)\right]$$

$$\text{or } \int_0^\infty d\omega\left[\frac{1}{\omega}\sin(\omega/2)\right] = \frac{\pi}{2}$$

$$\Rightarrow \int_0^\infty d\omega\left[\frac{1}{x}\sin(x/2)\right] = \frac{\pi}{2}$$

Fourier transform of trigonometric function, Gaussian, finite wave train & other functions

1. Find the Fourier transform of the function

$$f(x) = \begin{cases} b \ \ if \ \ |x| < a, \\ 0 \ \ if \ \ |x| > a \end{cases}$$

Solution : The F.T. of f(x) is given by

$$F(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx$$

$$= \frac{b}{\sqrt{2\pi}}\int_{-a}^{a} e^{-ikx}dx$$

$$= \frac{2ab}{\sqrt{2\pi}}\left\{\frac{\sin ka}{ka}\right\}$$

After obtaining $F(k)$ we can write

$$f(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} F(k)e^{ikx}dk$$

$$= \frac{1}{\sqrt{2\pi}}\frac{2b}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\frac{\sin ka}{k}\right\}e^{ikx}dk$$

$$= \frac{b}{\pi}\int_{-\infty}^{\infty}\left\{\frac{\sin ka}{k}\right\}e^{ikx}dk$$

2. Find the Fourier transform of the function

$$f(t) = \begin{cases} a, \ when \ -l < t < 0, \\ 0, \ otherwise, \ [a > 0] \end{cases}$$

Solution : The F. T. of f(t) is given by

$$f\,[f(t)] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$$

$$= \frac{a}{\sqrt{2\pi}}\int_{-l}^{0} e^{-i\omega t}dt$$

$$= \frac{2a}{\omega\sqrt{2\pi}}\sin(\omega l)$$

$$= \frac{ia}{\omega\sqrt{2\pi}}\left[1-e^{i\omega t}\right]$$

3. Find the Fourier transform of the function

$$f(t) = \begin{cases} a, \ when \ -l < t < 0, \\ 0, \ otherwise, \ [a > 0] \end{cases}$$

Solution : The F.T. of f(t) is given by

$$F[f(t)] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$$

$$= \frac{a}{\sqrt{2\pi}}\int_{-l}^{l} e^{-i\omega t}dt$$

$$= \frac{-a}{i\omega\sqrt{2\pi}}\left[e^{-i\omega t}\right]_{-l}^{l}$$

$$= \frac{-a}{i\omega\sqrt{2\pi}}\left[e^{i\omega l} - e^{-i\omega l}\right]$$

$$= \frac{2a}{\omega\sqrt{2\pi}}\left[\frac{e^{i\omega l} - e^{-i\omega l}}{2i}\right]$$

$$= \frac{2a}{\omega\sqrt{2\pi}}\sin(\omega l)$$

4. Find the Fourier transform of the function

$$f(x) = \begin{cases} 1, \ when \ |x| < l, \\ 0, \ when, \ |x| > l \end{cases}$$

Solution : The F.T. of f(x) is given by

$$F[f(x)] = f(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-i\omega x}dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-l}^{l} e^{-i\omega x}dx$$

$$= \frac{-1}{i\omega\sqrt{2\pi}}\left[e^{-i\omega x}\right]_{-l}^{l}$$

$$= \frac{1}{i\omega\sqrt{2\pi}}\left[e^{i\omega l} - e^{-i\omega l}\right]$$

$$= \frac{2}{\omega\sqrt{2\pi}}\left[\frac{e^{i\omega l} - e^{-i\omega l}}{2i}\right]$$

$$= \sqrt{\frac{2}{\pi}}\frac{\sin(\omega l)}{\omega}$$

For $\omega \to 0,\ F(\omega) = l\sqrt{\frac{2}{\pi}}$

$$f(t) = \begin{cases} \sin(\alpha t), & when\ 0 \le t \le \frac{\pi}{\alpha} \\ 0, & when,\ t > \frac{\pi}{\alpha} \end{cases}$$

Solution : Fourier sine transform of $A(k)$ of the function $f(t)$ is given by

$$A(k) = \sqrt{\frac{2}{\pi}}\int_{0}^{\pi/\alpha} f(t)\sin(kt)dt + \sqrt{\frac{2}{\pi}}\int_{\pi/\alpha}^{\infty} f(t)\sin(kt)dt$$

$$= \sqrt{\frac{2}{\pi}}\int_{0}^{\pi/\alpha} \sin(\alpha t)\sin(kt)dt$$

$$= \frac{1}{2}\sqrt{\frac{2}{\pi}}\int_{0}^{\pi/\alpha}\left[\cos(\alpha - k)t - \cos(\alpha + k)t\right]dt$$

$$= \frac{1}{2}\sqrt{\frac{2}{\pi}}\left[\frac{\sin\left[(\alpha-k)t\right]}{\alpha-k}-\frac{\sin\left[(\alpha+k)t\right]}{\alpha+k}\right]_0^{\pi/\alpha}$$

$$= \frac{1}{2}\sqrt{\frac{2}{\pi}}\left[\frac{\sin\left[(\alpha-k)\pi/\alpha\right]}{\alpha-k}-\frac{\sin\left[(\alpha+k)\pi/a\right]}{\alpha+k}\right]$$

(2.19)

6. Find the Fourier transform of the function $f(t)=e^{-\alpha|t|},\ -\infty< t<\infty,\ a>0$. write the inverse transform.

$$f(t)=\begin{cases} e^{at}, & t<0 \\ e^{at}, & t>0 \end{cases}$$

Solution : The F.T. of f(x) is given by

$$F\left[f(t)\right]=F(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)e^{i\omega t}dt$$

$$= \frac{1}{\sqrt{2\pi}}\left[\int_{-\infty}^{\infty} e^{at}e^{-i\omega t}dt+\int_0^{\infty} e^{-at}e^{-i\omega t}dt\right]$$

$$= \frac{1}{\sqrt{2\pi}}\left\{\left[\frac{e^{(a-i\omega)t}}{a-i\omega}\right]_{-\infty}^{0}+\left[\frac{e^{-(a+i\omega)t}}{a+i\omega}\right]_0^{\infty}\right\}$$

$$= \frac{1}{\sqrt{2\pi}}\left[\frac{1}{a-i\omega}+\frac{1}{a+i\omega}\right]$$

$$= \frac{1}{\sqrt{2\pi}}\left[\frac{2a}{a^2+\omega^2}\right]$$

$$= \sqrt{\frac{2}{\pi}}\frac{a}{\left(a^2+\omega^2\right)}$$

The inverse transform is given by,

$$\Phi(t) = F^{-1}\left[F(\omega)\right] = F^{-1}\left[\frac{a}{a^2+\omega^2}\right] = \sqrt{\frac{\pi}{2}}e^{-a|t|}.$$

2.4 Representation of Dirac delta function as a Fourier Integral

Dirac delta function

The one-dimensional Dirac delta function, $\delta(x)$, can be pictured as an infinitely high, infinitesimally narrow "spike," with unit area (i.e. area = 1 (as shown in figure 2.1 & 2.2).

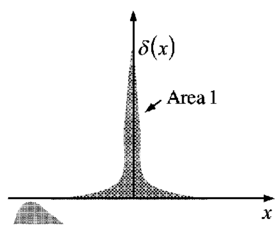


Figure 2.2 : Delta function with spike at x = 0

%

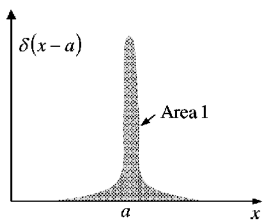


Figure 2.3 : Delta function with spike at x = a

That is to say :

$$\delta(x) = \begin{cases} 0, & if\ x \neq 0 \\ \infty, & if\ x = 0 \end{cases} \tag{2.20}$$

and

$$\int_{-\infty}^{\infty} \delta(x)dx = 1 \tag{2.21}$$

if $f(x)$ is some "ordinary" function then the product

$$f(x)\delta(x) = f(0)\delta(x) \tag{2.22}$$

Of course, we can shift the spike from x = o to some other point, $x = a$;

$$\delta(x-a) = \begin{cases} 0, & if\ x \neq a \\ \infty, & if\ x = a \end{cases} \tag{2.23}$$

and

$$\int_{-\infty}^{\infty} \delta(x-a)\,dx = 1. \qquad (2.24)$$

Also,

$$f(x)\delta(x-a) = f(a)\delta(x-a) \qquad (2.25)$$

and

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\,dx = f(a) \qquad (2.26)$$

1D representations of Dirac delta function :

$$\delta(x) = 0, \quad x \neq 0 \qquad (2.27)$$

$$\int_{-\infty}^{\infty} \delta(x)\,dx = 1 \qquad (2.28)$$

$$\int_{-\infty}^{\infty} f(x)\delta(x-x_0)\,dx = f(x_0) \qquad (2.29)$$

3D representations of Dirac delta function :

It is an easy matter to generalize the delta function to three dimensions :

$$\delta^3(\vec{r}) = \delta(x)\delta(y)\delta(z) \qquad (2.30)$$

Since, $\vec{r} = \hat{i}x + \hat{j}y + \hat{k}z$ is the position vector,extending from the origin to the point (x, y, z). This three-dimensional delta function is zero everywhere except at (0, 0, 0), where it blows up. Its volume integral is 1

$$\int_{All space} \delta^3(\vec{r})\,d\tau = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x)\delta(y)\delta(z)\,dxdydz = 1 \qquad (2.31)$$

and

$$\int_{All space} f(\vec{r})\delta^3(\vec{r}-\vec{a})\,d\tau = f(\vec{a} \qquad (2.32)$$

Since the divergence of $\frac{\hat{r}}{r^2}$ is zero everywhere except at the origin, and yet its integral over any volume containing the origin is a constant (4π). These are precisely the determining conditions for the Dirac delta function; evidently

$$\vec{\nabla}.\left(\frac{\hat{r}}{r^2}\right) = 4\pi\delta^3(\vec{r}) \tag{2.33}$$

Fourier integral representation of Dirac delta function

$$F(s) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\delta(x-x_0)e^{-isx}dx \tag{2.34}$$

$$= \frac{e^{isx_0}}{\sqrt{2\pi}} \tag{2.35}$$

$$F(s) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left[\delta(x+x_0)+\delta(x-x_0)\right]e^{-isx}dx \tag{2.36}$$

$$= \left[\frac{e^{-isx_0}+e^{isx_0}}{\sqrt{2\pi}}\right]$$

$$= \frac{2\cos(sx_0)}{\sqrt{2\pi}} \tag{2.37}$$

Solved examples

1. Evaluate the integral $I = \int_0^2 x^4\delta(x-3)$.

   Solution : As the spike of the function lies outside the domain of integration, the result of the integration would be 0.

2. Evaluate the integral $I = \int_0^5 x^4\delta(x-3)$.

   Solution: The delta function picks out the value of x^4 at the point $x = 3$, so the integral is $I = \int_0^5 x^4\delta(x-3) = f(3) = 3^4 = 81$.

3. Evaluate the integral $J = \int_{v}(r+1)\vec{\nabla}.\left(\frac{\hat{r}}{r^2}\right)d\tau$, where v is a sphere of radius R centered at the origin.

Solution : $J = \int_{\upsilon}(r+1)\vec{\nabla}.\left(\frac{\hat{r}}{r^2}\right)d\tau = \int_{\upsilon}(r+1)4\pi\delta^3(\vec{r})d\tau = 4\pi(0+1) = 4\pi$

2.5 Fourier transform of derivatives

Theorem 2.5.1 *Let for a positive integer n the n^{th} order derivative of $f(x)$, i,e., $f^{(n)}(x)$* is piecewise continuous on every interval $(-L, L)$ *and* $\int_{-\infty}^{\infty}\left|f^{(n-1)}(x)dx\right|$ *converges. Then one may asume that*

$$\lim_{x\to-\infty} f^{(k)}(x) = \lim_{x\to-\infty} f^{(k)}(x) = 0, 1, 2, ..., (n-1). \tag{2.38}$$

If $F\left[f'(x)\right] = F(\omega)$, *then*

$$F\left[f^{(n)}(x)\right] = (i\omega)^n F(\omega) \tag{2.39}$$

Proof : From the definition, we have

$$F\left[f'(x)\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f'(x)dx$$

$$= \frac{1}{\sqrt{2\pi}}\left[\left\{f(x)e^{-i\omega x}\right\}_{-\infty}^{\infty} + i\omega\int_{-\infty}^{\infty} f(x)e^{-i\omega x}dx\right]$$

$$= i\omega F(\omega)$$

$$F\left[f'(x)\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f''(x)dx$$

$$= \frac{1}{\sqrt{2\pi}}\left[\left\{f'(x)e^{-i\omega x}\right\}_{-\infty}^{\infty} + i\omega\int_{-\infty}^{\infty} f'(x)e^{-i\omega x}dx\right]$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f'(x)e^{-i\omega x}dx$$

$$= (i\omega)^2 F(\omega)$$

$= -\omega^2 F(\omega)$

By induction, we obtain the result

$$F\left[f^{(n)}(x)\right] = (i\omega)^n F(\omega)$$

2.6 Inverse Fourier transform

Fourier inversion theorem

The theorem states that, if $f(x)$ is square-integrable, then the Fourier transform $F\left[f(x)\right] = \hat{f}(k)$ exists and moreover

$$\int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk = \begin{cases} f(x), \ \text{if } f(x) \text{ is continuoius} \\ \frac{1}{2}\left[\lim_{y\to x+} + \lim_{y\to x-}\right] f(y), \ \text{otherwise} \end{cases} \tag{2.40}$$

In other words, at a point of discontinuity, the inverse transform produces the average of the and right limiting values of the function $f(x)$.

In Quantum mechanics (QM), Fourier transforms are defined slightly differently as

$$\Psi(k) = F\left[\psi(x)\right] = k\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x) e^{-\frac{ipx}{\hbar}} dx \tag{2.41}$$

and

$$\Psi(x) = F^{-1}\left[\Psi(k)\right] = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \Psi(k) e^{\frac{ipx}{\hbar}} dk \tag{2.42}$$

Applications

Solution of First order differential equation by Fourier Transform

1. Find the solution of the differential equation $y' - 2y = u_0(x)e^{-2x}, -\infty < x < \infty$ using Fourier transform.

   Solution : Applying FT on both sides of the gieven differential equation, we get

$$\mathcal{F}[y'] - 2\mathcal{F}[y] = \mathcal{F}[u_0(x)e^{-2x}]$$

$$(i\omega)Y(\omega) - 2Y(\omega) = \frac{1}{\sqrt{2\pi}}\frac{1}{(2+i\omega)}$$

$$Y(\omega) = \frac{1}{\sqrt{2\pi}}\frac{1}{(2+i\omega)(i\omega-2)}$$

$$= -\frac{1}{\sqrt{2\pi}}\frac{1}{(4+\omega^2)}$$

Where $\mathcal{F}[y(x)] = Y(\omega)$.
Using the formula for inverse FT :

$$\mathcal{F}^{-1}\left[\frac{1}{a^2+\omega^2}\right] = \frac{1}{a}\sqrt{\frac{\pi}{2}}e^{-a|t|}$$ when $a = 2$, we have:

$$y(x) = \mathcal{F}^{-1}[Y(\omega)]$$

$$= \mathcal{F}^{-1}\left[-\frac{1}{\sqrt{2\pi}}\frac{1}{4+\omega^2}\right]$$

$$= -\frac{1}{\sqrt{2}}\frac{1}{2}\sqrt{\frac{\pi}{2}}e^{-2|x|}$$

$$= -\frac{1}{4}e^{-2|x|}.$$

The solution cal also be expressed as

$$y(x) = \begin{cases} -\frac{1}{4}e^{2x}, & x < 0, \\ -\frac{1}{4}e^{-2x}, & x \geq 0. \end{cases}$$

2.7 Convolution theorem

Introduction

In mathematics, the convolution theorem states that under suitable conditions the Fourier transform of a convolution (meaning coil or winding or fold or twist) of two signals is the pointwise product of their Fourier transforms. This means convolution in one domain (e.g., time domain) equals point-wise multiplication in the other domain (e.g., frequency domain). The theorem holds for different Fourier-related transforms.

Theorem 2.7.1 *If f and g be two functions with convolution f * g (where asterisk denotes convolution in this context and not standard multiplication), then*

$$\mathcal{F}[f * g] = \mathcal{F}[f] * \mathcal{F}[g] \tag{2.43}$$

$$\mathcal{F}[f \otimes g] = \mathcal{F}[f] \otimes \mathcal{F}[g] \tag{2.44}$$

The tensor product symbol $\otimes$ is also used sometimes to represent convolution.

Definition of convolution

The convolution of piecewise continuous functions f, g is the function $f * g$ given by

$$(f * g)(t) = \int_0^t f(\tau) g(t - \tau) d\tau \tag{2.45}$$

Mathematical representation of convolution

$$c(u) = f(x) \otimes g(x) = \int_{-\infty}^{\infty} f(x) g(u - x) dx \tag{2.46}$$

Remarks :

1. $f * g$ is also called the generalized product of f and g.
2. The definition of convolution of two functions also holds in the case that one of the functions is a generalized function, like Dirac's delta.

Example of convolution of two functions

1. Find the convolution of $f(t) = e^{-t}$ and $g(t) = \sin(t)$.

   Solution : By definition:

$$\begin{aligned}
(f * g)(t) &= \int_0^t f(\tau) g(t - \tau) d\tau \\
&= \int_0^t e^{-\tau} \sin(t - \tau) d\tau \\
&= \left\{e^{-\tau} \cos(t - \tau)\right\}_0^t - \left\{e^{-\tau} \sin(t - \tau)\right\}_0^t - \int_0^t e^{-\tau} \sin(t - \tau) d\tau
\end{aligned}$$

$$\begin{aligned}
2\int_0^t e^{-\tau} \sin(t - \tau) d\tau &= \left[e^{-\tau} \cos(t - \tau)\right]_{\tau=0}^t - \left[e^{-\tau} \sin(t - \tau)\right]_{\tau=0}^t \\
&= e^{-t} - \cos t + \sin t
\end{aligned}$$

2.8 Properties of Fourier transforms: translation, change of scale, complex conjugation, etc

Translation (shifting)

Theorem 2.8.1 *If* $F(s)$ *is the complex Fourier transform of* $f(x)$*, then*

$$[A]: F\{f(x-a)\} = e^{isa}F(s) \quad (2.47)$$

$$[B]: F\left\{e^{iax}f(x)\right\} = F(s+a) \quad (2.48)$$

Proof : [A] :

$$F(s) = \int_{-\infty}^{\infty} e^{isx} f(x)dx$$

$$F\{f(x-a)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{isx} f(x-a); \; [put : x-a=t, \text{ so that } dx = dt]$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{is(t+a)} f(t)dt$$

$$g(x) = \frac{1}{2\sqrt{\pi\sigma t}}\int_{-\infty}^{\infty}$$

$$= \frac{e^{isa}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{ist} f(t)dt$$

$$= e^{isa}F(s)$$

Proof : [B] :

$$F\left\{e^{iax}f(x)\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{iax} f(x)e^{iex}dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i(s+a)x} f(x)dx$$

$$= F(s+a)$$

Change of Scale

Theorem 2.8.2 *If* $F(s)$ *is the complex Fourier transform of* $f(x)$*, then*

$$F\{f(ax)\} = \frac{1}{a}F\left(\frac{s}{a}\right)$$

Proof :

$$F\{f(ax)\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{isx} f(ax)dx; \left[Put : ax = t,\ dx = \frac{dt}{a}\right]$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{is\frac{t}{a}} f(t)\frac{dt}{a}$$

$$= \frac{1}{a}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\frac{s}{a}t} f(t)dt$$

$$= \frac{1}{a}F\left(\frac{s}{a}\right)$$

2.9 Three-dimensional Fourier transforms with examples

The 3-dimensioal Fourier transform

The Fourier pairs naturally extend to 3-dimensional functions as

$$\left.\begin{aligned} f(x,y,z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dk_x dk_y dk_z A\left(k_x,k_y,k_z\right)e^{i\left(k_x x+k_y y+k_z z\right)} \\ A\left(k_x,k_y,k_z\right) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdydz f(x,y,z)e^{-i\left(k_x x+k_y y+k_z z\right)} \end{aligned}\right\}$$

(2.49)

2.10 Application of Fourier Transforms to differential equations : One dimensional Wave adn Diffusion/Heat Flow Equations

One dimensional heat flow equations

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, [x, t > 0]$$

subject to the conditions :

1. $u = 0,\ when x = 0,\ t > 0$

2. $u = \begin{cases} 1, 0 \le x \le 1 \\ 0, x > 1 \end{cases}$, $when,\ t = 0$

3. $u(x, t)$ is bounded.

Note : If at x = 0, u is given then take Fourier sine transform and when $\frac{\partial u}{\partial x}$ at x = 0 is given, then use Fourier sine transform

For the given initial conditions : taking Fourier sine transform of both sides of given equation, we get

$$\int_{-\infty}^{\infty} \frac{\partial u}{\partial t} \sin sx dx = \int_0^{\infty} \frac{\partial^2 u}{\partial x^2} \sin sx dx \left[\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}\right]$$

$$\frac{\partial}{\partial t}\int_0^{\infty} u \sin sx dx = -s^2 \overline{u}(s) + su(0);\ [u = 0,\ when, x = 0]$$

$$\frac{\partial \overline{u}}{\partial t} = -s^2 \overline{u}(s)$$

$$\frac{\partial \overline{u}}{\partial t} + s^2 \overline{u}(s) = 0$$

$$(D + s^2)\overline{u} = 0$$

$$Let, \overline{u} = Ae^{mt}, then,$$

$$(m + s^2) = 0, [Auxiliary\ \ equation]$$

$$\Rightarrow m = -s^2 \tag{2.51}$$

$$\overline{u}(s,t) = Ae^{-s^2 t} \tag{2.52}$$

$$\overline{u} = \overline{u}(s,t) = \int_0^{\infty} u(x,t) \sin sx dx$$

$$\overline{u}(s,0) = \int_0^{\infty} u(x,t) \sin sx dx + \int_0^{\infty} u(x,0) \sin sx dx$$

$$= \int_0^{\infty} (1)\sin sxdx + \int_0^{\infty} (0)\sin sxdx$$

$$= \int_0^{\infty} (1)\sin sxdx$$

$$= \left[\frac{\cos sx}{s}\right]_0^1 = \frac{1-\cos s}{s} \quad (2.53)$$

$$But, \bar{u}(s,0) = A = \frac{1,\cos s}{x}, \left[From(2.65)\&(2.66)\right]$$

$$\bar{u}(s,t) = \left(\frac{1-\cos s}{s}\right)e^{-s^2 t} \quad (2.54)$$

Appling inverse FT on (2.67), we get

$$u(x,\ t) = \frac{2}{\pi}\int_0^{\infty} \bar{u}(s,t)\sin sxds \quad (2.55)$$

$$= \frac{2}{\pi}\int_0^{\infty} \left(\frac{1-\cos s}{s}\right)e^{-s^2 t}\sin sxds \quad (2.56)$$

One dimensional heat conduction

Consider a very long hetr conducting rod lying along the x-axis. The temperature-distribution in the rod is described by the one-dimensional heat equation:

$$\frac{\partial^2 \Theta}{\partial x^2} = \frac{1}{\sigma}\frac{\partial \Theta}{\partial t}\left[\Theta = \Theta(x,t=0);\, \Theta(x,t=0) = f(x)\right]$$

Asuuming thea, in the limit $x \to \pm\infty$, $\Theta(x,t) \to 0$, $\frac{\partial \Theta}{\partial t} \to 0$ and the Fourier transform of $f(x)$ exists. Let

$$\Theta(k,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \Theta(x,t)e^{ikx}dx$$

Taking Fourier transform of Eq. (2.60), we get

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{\partial^2 \Theta}{\partial x^2}e^{ikx}dx = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}}\frac{\partial}{\partial t}\int_{-\infty}^{\infty} \Theta(x,t)e^{ikx}dx.$$

Using trasforms of deriveatives, Eq. (2.62) becomes

$$-k^2\Theta(k,t) = \frac{1}{\sigma}\frac{\partial\Theta(k,t)}{\partial t} \tag{2.60}$$

where we assumed that, as $x \to \pm\infty$, $\Theta(x,t) \to 0$, $\frac{\partial\Theta}{\partial t} \to 0$. (i.e., the temperature is zero at the ends of a very long rod). The solution of Eq. (2.63) is

$$\Theta(k,t) = Ae^{-\sigma k^2 t} \tag{2.61}$$

Since

$$\Theta(k,0) = A = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\Theta(x,0)e^{ikx}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(x)e^{ikx}dx = f(k), \tag{2.62}$$

the solution in k-space becomes

$$\Theta(k,t) = F(k)e^{-\sigma k^2 t}. \tag{2.63}$$

The objective here is to derive an expression for $\Theta(x, t)$ for which need the inverse FT of $\Theta(k, t)$. This can be obtained by the use of the convolution theorem i.e.,

$$\Theta(x,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\left[F(k)e^{-\sigma k^2 t}\right]e^{-ikx}dk = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(\phi)g(x-\phi)d\phi \tag{2.64}$$

We require the original form (in x-space) of the second founction $g(x - \phi)$, which is just FT of $e^{-\sigma k^2 t}$ and is given by

$$g(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\sigma k^2 t}e^{-ikx}dk \tag{2.65}$$

Let $w^2 = 2k^2\sigma t,\ dk = \frac{dw}{\sqrt{2\sigma t}}$, then

$$g(x) = \frac{1}{2\sqrt{\pi\sigma t}}\int_{-\infty}^{\infty}e^{-\frac{w^2}{2}}e^{-i\left(\frac{w}{\sqrt{2\sigma t}}\right)x}dw$$

$$= \frac{1}{\sqrt{2\sigma t}}e^{-\frac{x^2}{4\sigma t}}$$

Using Eq. (2.70) for $(x-\phi)$, the expression for $\Theta(x,t)$ in terms of the convolution integral now becomes

$$\Theta(x,t)=\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\sigma t}}\int_{-\infty}^{\infty} f(\phi)e^{-\frac{(x-\phi)^2}{4\sigma t}}d\phi \tag{2.67}$$

The above Eq. (2.65) can be solved only when the specific form of $f(\phi)$ or $f(x)$, the initial temperature distribution along x are given.

$$\boxed{x=y} \tag{2.68}$$

Forced oscillation

Consider a system governed by a differential equation

$$\frac{d^2\phi(t)}{dt^2}+a_1\frac{d\phi(t)}{dt}+a_0\phi(t)=e^{i\omega t} \tag{2.69}$$

The function $\phi(t)$ represents the response of the system which is being driven by a sinusoidal force $e^{i\omega t}$. After considerably long time, a realistic system will be in a so-called steady state: in which

$$\phi(t)=A(\omega)e^{i\omega t}. \tag{2.70}$$

The reason is that the system dissipates energy due to damping or friction, so that in the absences of the driving term, the system will tend to lose all its energy: so that $(\phi)t\rightarrow 0$ in the limit as $t\rightarrow\infty$. To get the steady-state response of such systems one then substitutes $\phi(t)=A(\omega)e^{i\omega t}$. in the above equation (Eq. (2.37)) and solves for $A(\omega)$:

$$\frac{d^2\phi(t)}{dt^2}+a_1\frac{d\phi(t)}{dt}+a_0\phi(t)=\left(-\omega^2+a_1 i\omega+a_0\right)e^{i\omega t}=e^{i\omega t}$$

$$\Rightarrow A(\omega)=\frac{1}{-\omega^2+a_1 i\omega+a_0}$$

Therefore

$$\phi(t)=\frac{1}{-\omega^2+a_1 i\omega+a_0}e^{i\omega t} \tag{2.71}$$

Damped harmonic oscillator

the differential equation of motion of a damped harmonic oscillator is

$$\frac{d^2x(t)}{dt^2}+2b\frac{dx(t)}{dt}+\omega_0^2 s(t)=f(t) \tag{2.72}$$

where b is the damping paramater and $f(t)$ is any function of time such that

$$f(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}F(\omega)e^{-i\omega t}d\omega \tag{2.73}$$

and

$$F(\omega)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(t)e^{i\omega t}dt \tag{2.74}$$

Taking FT of each term in Eq. (2.56), we get

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{d^2x}{dt^2}e^{i\omega t}dt+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}2b\frac{dx}{dt}e^{i\omega t}dt$$

or, $$+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\omega_0^2 x(t)e^{i\omega t}dt=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}f(t)e^{i\omega t}dt$$

or, $$-\omega^2 X(\omega)+2bi\omega X(\omega)+\omega_0^2 X(w)=F(\omega)$$

or, $$X(\omega)=\frac{F(\omega)}{-\omega^2+2bi\omega+\omega_0^2}$$

or, $$X(\omega)=\frac{F(\omega)}{\omega_0^2-\omega^2+i2b\omega} \tag{2.75}$$

Again, $$X(\omega)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}x(t)e^{i\omega t}dt \tag{2.76}$$

Taking inverse FT of Eq. (2.60) we have,

$$x(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} X(\omega)e^{-i\omega t}dw \quad (2.77)$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{F(\omega)e^{-i\omega t}}{\omega_0^2-\omega^2+i2b\omega}d\omega \quad (2.78)$$

Chapter-end exercise

1. Find the fourier series of

   (a) $f(x)=x;\ 0<x>2\pi$.

   (b) $f(x)=-x;\ 0<x>2\pi$

   (c) $f(x)=-\frac{1}{2}x^2;\ -\pi<x>\pi$

2. Find the Fourier transform of

$$f(p)=\begin{cases}k, & |k|\le n\\ 0, & |k|>n\end{cases}$$

   where n is a positive integer.

3. Express the function

$$f(x)=x\sin x,\ -\pi,-\pi<x<\pi$$

   in the form of Fourier series and show that

$$\frac{\pi}{4}=\frac{1}{2}+\frac{1}{1.3}-\frac{1}{3.5}+\frac{1}{5.7}\cdots\cdots$$

[Ans: $f(x) = x\sin x = 1 - \frac{1}{2}\cos x - \frac{2}{1.3}\cos 2x + \frac{2}{2.4}\cos 3x - \frac{2}{3.5}\cos 4x + \frac{2}{4.6}\cos 5x - \frac{2}{5.7}\cos 6x$]

4. Show that if $F(-t) = -F(t)$, then real Fourier series expansion of $F(t)$ contains no sine terms.

5. Show that if $F(-t) = -F(t)$, then real Fourier series expansion of $F(t)$ contains no cosine terms and no constant term.

Unit 3 ❑ Laplace Transforms

Structure

3.1 Objectives

3.2 Introduction

3.3 Laplace transform

3.4 Laplace Transform (LT) of elementary functions

3.5 Properties of LTs: Change of scale theorem, Shifting theorem

3.6 LTs of 1st and 2nd order derivatives and integrals

3.7 Derivatives and integrals of LTs

3.8 LT of Unit Step function

3.9 Convolution Theorem

3.10 Inverse Laplace Transform (ILT)

3.11 Application of LT

3.1 Objective

After studying this section, students will be able to—

- This chapter aims to introduce the concept of Laplace transform and its inverse operations useful in solving linear as well as nonlinear differential equations, initial value problems, etc. Prerequisites for this chapter are the knowledge of integration and differentiation of functions of one and more variables.
- Successful learners will be able to apply this transform to solve the differential equation involving damped, undamped free, or forced oscillation as well as current-voltage related equations arising from simple electrical circuits containing inductor, resistor, a capacitor with constant or time-varying sources.

3.2 Introduction

In elementary calculus the students learned that **differentiation and integration are transforms**, meaning, these operations transform a function into another function.

For example, the function $f(t) = t^2$ is transformed, in turn, into a linear function and a family of cubic polynomial functions by the operations of differentiation and integration:

$$\frac{d}{dt}t^2 = 2t$$

and

$$\int t^2 dt = \frac{1}{3}t^3 + c\,.$$

Moreover, these two transforms possess the linearity property that the transform of a linear combination of functions is a linear combination of transforms. For a and b constants

$$\frac{d}{dt}\left[af(t)+bg(t)\right] = af'(t)+bg'(t)$$

and

$$\int\left[af(t)+bg(t)\right]dt = a\int f(t)dt + b\int g(t)dt$$

provided that each derivative and integral exists. In this section we shall examine a special type of integral transform called the **Laplace transform**. In addition to possessing the linearity property the Laplace transform has many interesting properties that make it very useful in solving linear initial-value problems.

Basic indea of integral transform and convergent, divergent integrals

If $f(x, y)$ is a function of two vaiables, then a definite integral of f with respect to one of the variables leads to a function of the other variable.

For example, by keeping x constant we see that

$$\int_1^2 2x^2 y dy = 3x^2$$

- Similarly, a definite integral such as

$$\int_a^b K(s,t)f(t)dt$$

transforms a function f of t into a function F of the variable s. We are particularly interested in an integral transfrom, where the interval of integration is the unbounded interval $[0,\infty]$. If $f(t)$ is defined for $t \geq 0$, then the improper integral is given as

$$\int_0^{\infty} K(s,t) f(t)dt = \lim_{b\to\infty} \int_0^b K(s,t) f(t)dt. \qquad (3.1)$$

If the limit in (3.1) exists, then we say that the integral exists or is convergent; if the limit does not exist, the integral does not exist and is divergent. The limit in (3.1) will, in general, exist for only certain values of the variable s.

Kernel

The function $K(s, t)$ in (3.1) is called the **kernel** of the transform. The choice $K(s, t) = e^{-st}$ as the kernel gives us a specially useful transform.

3.3 Laplace transform

Definition

If $f(t)$ be a function defined for all positive values of the variable t (i.e. $t \geq 0$), then the integral

$$\mathcal{L}\{f(t)\} = \int_0^{\infty} e^{-st} f(t)dt \qquad (3.2)$$

is said to be the Laplace transform (LT) of f, provided the integral conveges. This transform is named in the honor of the French mathematician astronomer Pierre-Simon Marquis de Laplace (1749-1827). When the defining integral (3.1) converges, the result is a finction of s. Generally we shall use a lower case letter to denote the function being transformed and the corresponding uppercase letter to denote its Laplace transform.

Notations used to denote Laplace transforms

$\mathcal{L}\{f(t)\} = F(s)$, $\mathcal{L}\{g(t)\} = G(s)$, $\mathcal{L}\{x(t)\} = X(s)$ or some times $\mathcal{L}[f(t)] = F(s)$

An example based on the definnition of Laplace transform

1. Evaluate L{1}.

   Solution: From definition (3.2),

$$\mathcal{L}\{1\} = \int_0^\infty (1)e^{-st}dt$$

$$= \lim_{b\to\infty}\int_0^b e^{-st}dt$$

$$= \lim_{b\to\infty}\left.\frac{-e^{-st}}{s}\right|_0^b$$

$$= \lim_{b\to\infty}\frac{-e^{-sb}+1}{s}$$

$$= \frac{1}{s},$$

provided that $s > 0$. In other words, when $s > 0$, the exponent $-sb$ is negative, and $e^{-sb} \to 0$ as $b \to \infty$. The integral diverges for $s < 0$.

Some important formulae on Laplace transforms and their proofs

$$\mathcal{L}[1] = \frac{1}{s}, s > 0 \quad (3.3)$$

$$\mathcal{L}[1] = \frac{n!}{s^{n+1}}, n = 0,1,2,3,.... \quad (3.4)$$

$$\mathcal{L}[e^{ct}] = \frac{1}{s-c};[s > c] \quad (3.5)$$

$$\mathcal{L}[e^{-ct}] = \frac{1}{s+c}; s,c > 0 \quad (3.6)$$

$$\mathcal{L}[\cos ct] = \frac{s}{s^2+c^2};[s > 0] \quad (3.7)$$

$$\mathcal{L}[\sin ct] = \frac{c}{s^2+c^2};[s > 0] \quad (3.8)$$

$$\mathcal{L}[\cosh ct] = \frac{c}{s^2-c^2};[s^2 > c^2] \quad (3.9)$$

$$\mathcal{L}[\sinh ct] = \frac{c}{s^2-c^2};[s^2 > c^2] \quad (3.10)$$

Proof of the formulae

$$\boxed{\mathcal{L}[1] = \frac{1}{s}}$$

$$L(1) = \int_0^\infty e^{-st}(1)dt = \left[\frac{e^{-st}}{-s}\right]_0^\infty = \frac{1}{-s}\left[e^{-\infty} - e^0\right] = \frac{1}{-s}[0-1] = \frac{1}{s}$$

$$\boxed{\mathcal{L}[t^n] = \frac{n!}{s^{n+1}}}$$

$$L(t^n) = \int_0^\infty e^{-st}(t^n)dt = \int_0^\infty e^{-x}\left(\frac{x^n}{s^n}\right)\left(\frac{dx}{s}\right) = \frac{1}{s^{n+1}}\int_0^\infty e^{-x}x^n dx = \frac{\Gamma(n+1)}{s^{n+1}} = \frac{n!}{s^{n+1}}$$

$$\boxed{\mathcal{L}[e^{ct}] = \frac{1}{s-c}}$$

$$L(e^{ct}) = \int_0^\infty e^{-st}(e^{ct})dt = \int_0^\infty e^{-(s-c)t}dt = \left[\frac{e^{-(s-c)t}}{-(s-c)}\right]_0^\infty = \frac{1}{-(s-c)}[e^{-\infty} - e^0] =$$

$$\frac{1}{-(s-c)}[0-1] = \frac{1}{s-c}$$

$$\boxed{\mathcal{L}\left[e^{-ct}\right] = \frac{1}{s+c}}$$

$$L(e^{-ct}) = \int_0^\infty e^{-st}(e^{-ct})dt = \int_0^\infty e^{-(s+c)t}dt = \left[\frac{e^{-(s+c)t}}{-(s+c)}\right]_0^\infty = \frac{1}{-(s+c)}[e^{-\infty} - e^0] =$$

$$\frac{1}{-(s+c)}[0-1] = \frac{1}{s+c}$$

$$\boxed{\mathcal{L}[\cos ct] = \frac{s}{s^2+c^2}}$$

$$L(\cos ct) = \int_0^\infty e^{-st}(\cos ct)dt = \int_0^\infty e^{-st}\left(\frac{e^{ict}+e^{-ict}}{2}\right)dt = \frac{1}{2}\left[\int_0^\infty e^{-(ic)t}dt + \right.$$

$$\left.\int_0^\infty e^{-(s+ic)t}dt\right] = \frac{1}{2}\left[\frac{1}{s-ic} + \frac{1}{s+ic}\right] = \frac{s}{s^2+c^2}$$

$$\boxed{\mathcal{L}[\sin ct] = \frac{c}{s^2+c^2}}$$

$$L(\sin ct) = \int_0^\infty e^{-st}(\sin ct)dt = \int_0^\infty e^{-st}\left(\frac{e^{ict}-e^{-ict}}{2i}\right)dt = \frac{1}{2i}\left[\int_0^\infty e^{-(s-ic)t}dt - \int_0^\infty e^{-(s+c)t}dt\right] = \frac{1}{2}\left[\frac{1}{s-c}+\frac{1}{s+c}\right] = \frac{s}{s^2-c^2}$$

$$\boxed{\mathcal{L}[\cosh ct] = \frac{s}{s^2-c^2}}$$

$$L(\cosh ct) = \int_0^\infty e^{-st}(\cos ct)dt = \int_0^\infty e^{-st}\left(\frac{e^{ct}+e^{-ct}}{2}\right)dt = \frac{1}{2}\left[\int_0^\infty e^{-(s-c)t}dt + \int_0^\infty e^{-(s+c)t}dt\right] = \frac{1}{2}\left[\frac{1}{s-c}+\frac{1}{s+c}\right] = \frac{s}{s^2-c^2}$$

$$\boxed{\mathcal{L}[\sinh ct] = \frac{c}{s^2-c^2}}$$

$$L(\sinh ct) = \int_0^\infty e^{-st}(\sinh ct)dt = \int_0^\infty e^{-st}\left(\frac{e^{ct}-e^{-ct}}{2}\right)dt = \frac{1}{2}\left[\int_0^\infty e^{-(s-c)t}dt - \int_0^\infty e^{-(s+c)t}dt\right] = \frac{1}{2}\left[\frac{1}{s-c}+\frac{1}{s+c}\right] = \frac{c}{s^2-c^2}$$

3.4 Laplace Transform (LT) of elementary functions

Solved examples on Laplace transform of elementary functions

1. Find the Laplace transform of the function $f(x) = \begin{cases} 0, f\ or\ x < 3 \\ x-3, f\ or\ x \geq 3. \end{cases}$

Solution:

$$\mathcal{L}[f(x)] = \int_0^\infty e^{-sx} f(x)dx$$
$$= \int_0^3 e^{-sx} f(x)dx + \int_3^\infty e^{-sx} f(x)dx$$
$$= \int_0^3 \left(e^{-sx}\right)(0)\,dx + \int_3^\infty \left(e^{-sx}\right)(x-3)\,dx$$
$$= \int_3^\infty e^{-sx}(x-3)dx$$
$$= \left[(x-3)\frac{e^{-sx}}{-s}\right]\Bigg|_3^\infty - \int_3^\infty \left(\frac{e^{-sx}}{-s}\right)dx$$
$$= 0 - \left[\frac{e^{-sx}}{(-s)^2}\right]_{x=3}^{\infty}$$
$$= \frac{e^{-3s}}{s^2}.$$

2. Find the Laplace transform, if it exists, of each of the following functions

(a) $f(t) = e^{at}$

(b) $f(t) = 1$

(c) $f(t) = t$

Solution:

2. (a)

$$\mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}dt = \int_0^\infty e^{at}e^{-st}dt = \int_0^\infty e^{-(s-a)t}dt = \left[\frac{e^{-(s-a)t}}{-(s-a)}\right]_0^\infty =$$

$\frac{1}{s-a}$ for $s > a$.

2. (b)

$$\mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}dt = \int_0^\infty (1)e^{st}dt = \int_0^\infty e^{-st}dt = \left[\frac{e^{-st}}{-s}\right]_0^\infty = \frac{1}{s}$$

for $s > 0$.

2. (c)

$$\mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}dt = \int_0^\infty (t)e^{-st}dt = \int_0^\infty te^{-st}dt = \left[\frac{te^{-st}}{-s} + \frac{e^{-st}}{-s^2}\right]_0^\infty =$$

$\frac{1}{s^2}$ for $s > 0$.

3.5 Properties of LTs: Change of scale theorem, Shifting theorem

We shall continue discussing various properties of Laplace transform. We mainly cover change of scale property, Laplace transform of integrals and derivatives etc.

Change of scale property

$$\boxed{\mathcal{L}[f(ct)] = \frac{1}{c}F\left(\frac{s}{c}\right)}$$

Proof:

If $\mathcal{L}\{f(t)\} = F(s)$, then $\mathcal{L}\{f(ct)\} = \int_0^\infty e^{-st}[f(ct)]dt = \int_0^\infty e^{-(s/c)u}[f(u)]\left[\frac{du}{c}\right] =$

$\frac{1}{c}\int_0^\infty e^{qu} f(u)du = \frac{1}{c}F(q) = \frac{1}{c}F\left(\frac{s}{c}\right)$; [*Assumptions* : $ct = u$, $dt = du/c$, $q = s/c$]

Shifting theorem

$$\boxed{\mathcal{L}\{f(t-c)\} = e^{-cs}F(s)}$$

Proof:

If $\mathcal{L}\{f(t)\} = F(s)$, then $\mathcal{L}\{f(t-c)\} = \int_0^\infty e^{-st}[f(t-c)]dt = \int_0^\infty e^{-s(c+p)}[f(p)]dp =$

$e^{-sc}\int_0^\infty e^{-sp} f(p)dp = e^{sc}F(s)$; [*Assumptions* : $t-c = p, t = c+p, dt = dp$]

3.6 Laplace Transforms of first and second order derivatives and integrals of functions

LT of 1st order derivative of functions

Let $f(t)$ be a function and $f'(t)$ be its derivative with respect to time which exists. Laplace transform of $f(t)$ is given by

$$\mathcal{L}[f(t)] = F(s) = \int_0^\infty f(t)e^{-st}dt \tag{3.11}$$

Laplace transform of $f'(t)$ is

$$\mathcal{L}[f'(t)] = \int_0^\infty f'(t)e^{-st}dt \tag{3.12}$$

Carrying integration by parts of Eq. (3.12) we get

$$\begin{aligned}\mathcal{L}[f'(t)] &= \left[e^{-st}f(t)\right]_0^\infty - \int_0^\infty\left(-se^{-st}\right)f(t)dt\\ &= -f(0) + s\int_0^\infty f(t)e^{-st}dt\\ &= -f(0) + sF(s); \text{[using Eq. (3.11)]}\\ &= sF(s) - f(0)\end{aligned} \tag{3.13}$$

LT of 2nd order derivative of functions

Let $f'(t)$ and $f''(t)$ denote the 1st and 2nd order derivaties of the function $f(t)$. Laplace transform of $f''(t)$ is

$$\mathcal{L}[f''(t)] = \int_0^\infty f''(t)e^{-st}dt \tag{3.14}$$

Carrying integration by parts of Eq. (3.14) we get

$$\begin{aligned}\mathcal{L}[f''(t)] &= \left[e^{-st}f'(t)\right]_0^\infty - \int_0^\infty\left(-se^{-st}\right)f'(t)dt\\ &= -f'(0) + s\int_0^\infty f'(t)e^{-st}dt\\ &= -f'(0) + s\mathcal{L}[f'(t)]\\ &= -f'(0) + s[sF(s) - f(0)]; \text{[using Eq. (3.13)]}\\ &= s^2F(s) - sf(0) - f'(0)\end{aligned} \tag{3.15}$$

LTs of integrals of functions

Let us consider the function defined in terms of the integral

$$g(t) = \int_0^t f(\tau)d\tau \tag{3.16}$$

such that

$$g'(t) = f(t) \tag{3.17}$$

Now,

$$\begin{aligned} \mathcal{L}[f(t)] &= \mathcal{L}[g'(t)]; \text{[using Eq. (3.17)]} \\ &= s\mathcal{L}[g(t)] - g(0); \text{[using Eq. (3.13)]} \end{aligned} \tag{3.18}$$

For $g(0) = 0$, we have from Eq. (3.18)

$$\begin{aligned} F(s) &= s\mathcal{L}[g(t)] \\ &= s\mathcal{L}\left[\int_0^t f(\tau)d\tau\right] \\ \text{or, } \mathcal{L}\left[\int_0^t f(\tau)d\tau\right] &= \frac{F(s)}{s} \end{aligned} \tag{3.19}$$

3.7 Derivatives and integrals of LTs

Derivatives of LT

If $L[F(t)] = f(s)$, then $f'(s) = \mathcal{L}[-tf(t)]$ and $\frac{d^n}{ds^n} f(s) = (-1)^n \mathcal{L}\left[t^n F(t)\right]$

Proof :

By definition,

$$f(s) = \int_0^\infty e^{-st} F(t)dt \tag{3.20}$$

Differentiating w.r.s.

$$f'(s) = \int_0^\infty -te^{-st} F(t)dt = \mathcal{L}[-tF(t)] \tag{3.21}$$

By repeated differentiations,

$$f^{n}(s)=\int_0^{\infty}(-t)^n e^{-st}F(t)dt=\mathcal{L}\left[(-t)^n F(t)\right]=(-1)^n\,\mathcal{L}\left[t^n F(t)\right] \qquad (3.22)$$

Integrals of LT

If $L[F(t)]=f(s)$, then $\mathcal{L}\left[\int_0^t F(p)dp\right]=\frac{f(s)}{s}$

Proof:

By definition,

$$\mathcal{L}\left[\int_0^t F(p)dp\right]=\int_0^{\infty}\left[\int_0^t F(p)dp\right]e^{-st}dt$$

$$=\left[\frac{e^{-st}}{s}\int_0^t F(p)dp\right]_0^{\infty}+\frac{1}{s}\int_0^{\infty}F(t)e^{-st}dt$$

Using integration by parts

$$=\frac{1}{s}L[F(t)]$$

$$=\frac{f(s)}{s}$$

For $F(t)$ at least piecewise continuos and x large enough so that $e^{-xt}F(t)$ decreases exponentially $(\text{as } x\to\infty)$, the integral

$$f(x)=\int_{t=0}^{\infty}e^{-xt}F(t)dt \qquad (3.23)$$

is uniformly convergent with respect to x. This justifies reversing the order of integration in the following equation:

$$\int_{x=s}^{\infty}f(x)dx=\int_{x=s}^{b}dx\int_{t=0}^{\infty}dte^{-xt}F(t) \qquad (3.24)$$

$$=\int_{t=0}^{\infty}\frac{F(t)}{t}\left(e^{-st}-e^{-bt}\right)dt, \qquad (3.25)$$

on integrating with respect to x. The lower limit is chosen large enough so that $f(s)$ is within the region of uniform convegence. Now letting $b\to\infty$, we have

$$\int_{x=s}^{\infty} f(x)dx = \int_0^{\infty} \frac{F(t)}{t} e^{-st} dt = L\left[\frac{F(t)}{t}\right], \qquad (3.26)$$

provided that $\frac{F(t)}{t}$ is finite at $t = 0$ or diverge less strongly than t^{-1} (so that $L\left[\frac{F(t)}{t}\right]$ will exist.

3.8 LT of Unit Step function, Dirac Delta function, Periodic Functions

LT of Unit Step function

$$\boxed{\mathcal{L}\left[u(t-c)\right] = \frac{1}{c}F(s)}$$

Q. Find the Laplace transform of the function un it step function defined as:

$$u(t-c) = \begin{cases} 0, f\text{ or } t < c \\ 1, f\text{ or } t \geq c. \end{cases}$$

Solution:

$$\mathcal{L}[u(t-c)] = \int_0^{\infty} e^{-st} u(t-c)dt$$

$$= \int_0^{c} e^{-st} u(t-c)dt + \int_c^{\infty} e^{-st} u(t-c)dt$$

$$= \int_0^{c} \left(e^{-st}\right)(0)dt + \int_3^{\infty} \left(e^{-st}\right)(1)dt$$

$$= \int_c^{\infty} e^{-st} dt$$

$$= \left[\frac{e^{-st}}{-s}\right]\Bigg|_c^{\infty}$$

$$= \left[\frac{0 - c^{-cs}}{-s}\right]$$

$$= \frac{e^{-cs}}{s}$$

LT of Dirac delta function

$$\boxed{\mathcal{L}\left[u\left(t-c\right)\right]=\frac{1}{c}F(s)}$$

Q. Find the Laplace transform of the unit step function defined as:

$$u\left(t-c\right)=\begin{cases}0, f\text{ or } t<c\\1, f\text{ or } t\geq c.\end{cases}$$

Solution:

$$\begin{aligned}\mathcal{L}\left[u(t-c)\right]&=\int_0^{\infty}e^{-st}u(t-c)dt\\&=\int_0^{c}e^{-st}u(t-c)dt+\int_c^{\infty}e^{-st}u\left(t-c\right)dt\\&=\int_0^{c}\left(e^{-st}\right)(0)dt+\int_3^{\infty}\left(e^{-st}\right)(1)dt\\&=\int_c^{\infty}e^{-st}dt\\&=\left[\frac{e^{-st}}{-s}\right]\Big|_c^{\infty}\\&=\left[\frac{0-c^{-cs}}{-s}\right]\\&=\frac{e^{-cs}}{s}\end{aligned}$$

LT of periodic function

$$\boxed{\mathcal{L}\left[f\left(t\right)\right]=\frac{\int_0^{T}e^{-st}f(t)dt}{1-e^{-st}}}$$

Q. Find the Laplace transform of the periodic function defined as:

$$f\left(t\right)=\begin{cases}t, & f\text{ or } 0<t\leq c\\2c-t, f\text{ or } c<t\leq 2c.\end{cases}$$

Solution: Period = $T = 2c$. Laplace transform of periodic function $f(t)$ is

$$\mathcal{L}[f(t)] = \frac{\int_0^T e^{-st} f(t)dt}{1-e^{-sT}} = \frac{\int_0^{2c} e^{-st} f(t)dt}{1-e^{-2cs}}.$$

On puting the values of $f(t)$, we get

$$\mathcal{L}[f(t)] = \frac{1}{1-e^{-2cs}}\left[\int_0^{2c} e^{-st}(c)dt + \int_c^{2c} e^{-st}(2c-t)dt\right]$$

$$= \frac{1}{1-e^{-2cs}}\left[\left\{\frac{te^{-st}}{-s} - \frac{e^{-st}}{(-s)^2}\right\}_0^c + \left\{(2c-t)\frac{e^{-st}}{(-s)} - (-1)\frac{e^{-st}}{(-s)^2}\right\}_c^{2c}\right]$$

$$= \frac{1}{1-e^{-2cs}}\left[\left\{\frac{ce^{-cs}}{(-s)} - \frac{e^{-cs}}{(-s)^2} - 0 + \frac{1}{(-s)^2}\right\}\right.$$

$$\left. + \left\{(2c-2c)\frac{e^{-2cs}}{(-s)} + \frac{e^{-2cs}}{s^2} - \left((2c-c)\frac{e^{-cs}}{(-s)}\right) + \frac{e^{-cs}}{s^2}\right\}\right]$$

$$= \frac{1}{1-e^{-2cs}}\left[-\frac{ce^{-cs}}{s} - \frac{c^{-cs}}{s^2} + \frac{1}{s^2} + \frac{e^{-2cs}}{s^2} + \frac{ce^{-cs}}{s} - \frac{e^{-cs}}{s^2}\right]$$

$$= \frac{1}{1-e^{-2cs}}\left[\frac{1}{s^2}\left(1-2e^{-cs}+e^{-2cs}\right)\right]$$

$$= \frac{1}{s^2}\left[\frac{(1-e^{-cs})^2}{(1+e^{-ce})(1-e^{-cs})}\right]$$

$$= \frac{1}{s^2}\left[\frac{1-e^{-cs}}{1+e^{-cs}}\right]$$

3.9 Convolution Theorem

If $f(t)$ and $g(t)$ be two functions of t. The convolution of $f(t)$ and $g(t)$ is also a function of t, denoted by $(f * g)(t)$ and is defined by the relation

$$(f*g)(t) = \int_{-\infty}^{\infty} f(t-x)g(x)dx \tag{3.27}$$

However if f and g are both causal functions then (strictly) $f(t)g(t)$ are written $f(t)u(t)$ and $g(t)u(t)$ respectively, so that

$$(f*g)(t)=\int_{-\infty}^{\infty} f(t-x)u(t-x)g(x)u(x)dx=\int_0^t f(t-x)g(x)dx \qquad (3.28)$$

because of the properties of the step functions ($u(t-x)=0$ if $x>t$ and $u(x)=0$ if $x<0$).

3.10 Inverse Laplace Transform (ILT)

Example of Inverse Laplace Transform

1. If $\mathcal{L}^{-1}\left[\frac{s}{s^2-16}\right]=\cosh 4t$ then determine $\mathcal{L}^{-1}\left[\frac{s}{2s^2-8}\right]$

Solution: Given that $\mathcal{L}^{-1}\left[\frac{s}{s^2-16}\right]=\cosh 4t$. Replacing s by $2s$ and using scaling property we have $\mathcal{L}^{-1}\left[\frac{2s}{4s^2-16}\right]=\frac{1}{2}\cosh 2t$ $\mathcal{L}^{-1}\left[\frac{s}{2s^2-8}\right]=\frac{1}{2}\cosh 2t$

2. Find $\mathcal{L}^{-1}\left[\frac{1}{s-a}\right], s>1$

3. Find $\mathcal{L}^{-1}\left[\frac{3}{s-2}\right]$

4. Find $\mathcal{L}^{-1}\left[\frac{2}{s+2}+\frac{2}{s-2}\right]$

Change of Scale Property

If $\mathcal{L}^{-1}[F(s)]=f(t)$, then $\mathcal{L}^{-1}[F(as)]=\frac{1}{a}F\left(\frac{t}{a}\right)$

Inverse Laplace Transform of derivatives (Derivative Theorem)

If $\mathcal{L}^{-1}[F(s)]=f(t)$, then $\mathcal{L}^{-1}\left[\frac{d^n}{ds^n}f(as)\right]=(-1)^n t^n f(t), n=1,2$

Examples

1. Find the ILT of (i) $\frac{2as}{\left(s^2+a^2\right)^2}$ (ii) $\frac{s^2-a^2}{(s^2+a^2)^2}$

Solution: Note that $\frac{d}{ds}\left[\frac{a}{s^2+a^2}\right]=-\frac{2as}{\left(s^2+a^2\right)^2}$ and $\frac{d}{ds}\left[\frac{s}{s^2+a^2}\right]=-\frac{s^2-a^2}{\left(s^2+a^2\right)^2}$

(i) By direct application of the derivative theorem we have,

$$L^{-1}\left[\frac{2as}{\left(s^2+a^2\right)^2}\right]=L^{-1}\left[\frac{d}{ds}\left[-\frac{a}{s^2+a^2}\right]\right]=(-1)t\,L^{-1}\left[-\frac{a}{s^2+a^2}\right]=t\sin at \text{ and}$$

(ii) $$L^{-1}\left[\frac{s^2-a^2}{\left(s^2+a^2\right)^2}\right]=L^{-1}\left[-\frac{d}{ds}\left[\frac{s}{s^2+a^2}\right]\right]=(-1)t\,L^{-1}\left[-\frac{s}{s^2+a^2}\right]=t\cos at$$

Inverse Laplace Transform of Integrals

If $\mathcal{L}^{-1}\left[F(s)\right]=f(t)$, *then* $\mathcal{L}^{-1}\left[\int_s^\infty f(s)\right]=\frac{f(t)}{t}$

1. Find the inverse Laplace transform (ILT) f(t) of the function

$$\int_s^\infty \frac{1}{s(s+1)}ds$$

Solution: Solution: By the method of partial fraction we obtain

$$\mathcal{L}^{-1}\left[\int_s^\infty \frac{1}{s(s+1)}ds\right]=\mathcal{L}^{-1}\left[\int_s^\infty \left\{\frac{1}{s}-\frac{1}{s+1}\right\}ds\right]$$

$$=\mathcal{L}^{-1}\left[\int_s^\infty \frac{1}{s}ds\right]-\mathcal{L}^{-1}\left[\int_s^\infty \frac{1}{s+1}ds\right]$$

$$=\frac{1-e^{-t}}{t}$$

Multiplication by powers of s

If $\mathcal{L}^{-1}[F(s)] = f(t)$, and $f(0) = 0$, then $\mathcal{L}^{-1}[sF(s)] = f'(t)$

1. Using $\mathcal{L}^{-1}\left[\frac{1}{s^2+1}\right] = \sin t$, and with the application of the above result evaluate

$\mathcal{L}^{-1}\left[\frac{s}{s^2+1}\right]$

Solution: Direct application of the above result leads to

$$\mathcal{L}^{-1}\left[\frac{s}{s^2+1}\right] = \frac{d}{dt}\sin t = \cos t$$

3.11 Application of LT to 2nd order differential equations

Driven Damped Harmonic Oscillator

Many systems exhibit mechanical stability: disturbed from an equilibrium position they move back toward that equilibrium position. We will start with a single degree of freedom, which will illustrate most of the important behavior. If the damped oscillator is driven by an arbitrary function of time, then

$$\ddot{x} + 2b\dot{x} + \omega_0^2 x = \frac{F(t)}{m}. \tag{3.29}$$

One of the important techniques for solving Eq. (3.10) is via an integral transform called the Laplace transform. The transform converts time derivatives into polynomials, which produces an algebraic equation. These are easier to solve than differential equations. After solving the algebraic equation, we then apply the inverse Laplace transform to return to a time-domain expression that gives x(t).

First, we define the Laplace transform of a function of time, $x(t)$, as

$$X(s) = \mathcal{L}[x(t)] = \int_0^\infty x(t)e^{-st}dt. \tag{3.30}$$

Since we will be applying this to time derivatives of generalized coordinates, let's work out expressions for the Laplace transform of $\frac{dx}{dt}$ and $\frac{d^2x}{dt^2}$.

$$\mathcal{L}\left[\frac{dx}{dt}\right] = \int_0^\infty \frac{dx}{dt} e^{-st} dt,\ [x = x(t)] \tag{3.31}$$

$$= e^{-st} x(t)\Big|_0^\infty + \int_0^\infty s e^{-st} x dt$$

$$= -x(0) + sX(s)$$

$$\mathcal{L}\left[\frac{dx}{dt}\right] = sX(s) - x(0) \tag{3.32}$$

$$\mathcal{L}\left[\frac{d^2x}{dt^2}\right] = \int_0^\infty \frac{d^2x}{dt^2} e^{-st} dt. \tag{3.33}$$

$$= e^{-st} \left.\frac{dx}{dt}\right)\Big|_0^\infty + \int_0^\infty s e^{-st} \frac{dx}{dt} dt$$

$$= -x'(0) + s(sX(s) - x(0))$$

$$\mathcal{L}\left[\frac{d^2x}{dt^2}\right] = s^2 X(s) - sx(0)) - x'(0) \tag{3.34}$$

where the prime indicates differentiation with respect to the argument of the function. We now seek to apply the Laplace transform to Eq. (3.10). Let us assume that we have situated the origin of time such that the system is quiescent and the forcing function vanishes for $t < 0$. Then we may take all of the integrated terms to vanish, and get

$$(s^2 + 2\beta s + \omega_0^2) X(s) = \mathcal{L}\left[\frac{F(t)}{m}\right] = \frac{F(s)}{m} \tag{3.35}$$

Solving for the Laplace transform of x, $X(s)$, gives

$$X(s) = \frac{F(s)}{m(s^2 + 2\beta s + \omega_0^2)} \tag{3.36}$$

We now apply the Laplace transform inversion integral l (the Bromwich integral),

$$x(t) = \frac{1}{2\pi i} \int_{\gamma-\infty}^{\gamma+\infty} X(s) e^{st} ds \tag{3.37}$$

$$= \frac{1}{2\pi i} \int_{\gamma-\infty}^{\gamma+\infty} \frac{F(s)}{m(s^2 + 2\beta s + \omega_0^2)} e^{st} ds \tag{3.38}$$

where γ is a real constant that exceeds the real part of all the singularities of $X(s)$, to solve for $x(t)$. Equation (3.19) is derived on the basis of contour integration.

Example 1

Suppose that the oscillator described in Eq. (3.10) is thumped at $t = 0$ with a delta function impulse: $F(t) = \alpha\delta(t)$. Find $x(t)$ using the Laplace transform method. According to Eq. (3.17), we need first to calculate the Laplace transform of the forcing function:

$$L\left[F(t)\right] = L\left[\alpha\delta(t)\right] = \int_0^{\infty} \alpha\delta(t)e^{-st}dt = \alpha \qquad (3.39)$$

By Eq. (3.17) and the inversion integral, we have

$$x(t) = \frac{1}{2\pi i}\int_{\gamma-\infty}^{\gamma+\infty} \frac{\alpha/m}{s^2 + 2\beta s + \omega_0^2} e^{st} ds \qquad (3.40)$$

The integrand has poles at the roots of the quadratic equation in the denominator, $s_{\pm} = -\beta \pm \sqrt{\beta^2 - \omega_0^2}$, both of which lie to the left of s = 0. So, we may integrate along the imaginary axis and close in the left half-plane, where the exponential sends the integrand to zero for t > 0. By the residue theorem, the integral is therefore $2\pi i$ times the sum of the two residues.

Writing the denominator as $(s - s_+)(s - s_-)$

$$x(t) = \frac{1}{2\pi i}\int_{\gamma-\infty}^{\gamma+\infty} \frac{\alpha/m}{(s-s_+)(s-s_-)} e^{st} ds$$

$$= \frac{1}{2\pi i}\int_{\gamma-\infty}^{\gamma+\infty} \frac{\alpha/m}{(s_+ - s_-)}\left[\frac{1}{s-s_+} - \frac{1}{s-s_-}\right] e^{st} ds$$

$$= \frac{1}{2\pi i}\frac{\alpha/m}{(s_+ - s_-)}(2\pi i)\left[\lim_{s\to s_+}\frac{(s-s_+)}{s-s_+}e^{st} - \lim_{s\to s_-}\frac{(s-s_-)}{s-s_-}e^{st}\right]$$

$$= \frac{\alpha/m}{(s_+ - s_-)}\left[e^{s+t} - e^{s-t}\right]$$

$$= \frac{\alpha}{m}\frac{1}{2}\frac{e^{-\beta t}}{\sqrt{\beta^2-\omega_0^2}}\left(e^{\sqrt{\beta^2-\omega_0^2}t} - e^{-\sqrt{\beta^2-\omega_0^2}t}\right)$$

$$= \frac{\alpha}{m}\frac{e^{-\beta t}}{\omega}\left(\frac{e^{i\omega t} - e^{-i\omega t}}{2i}\right) \qquad (3.41)$$

Finally, $x(t) = \frac{\alpha}{m}\frac{e^{-\beta t}}{\omega}\sin(\omega t)$

where $\omega = \sqrt{\omega_0^2 - \beta^2}$

Simple Electrical Circuits

LCR series circuit with constant voltage source

Let us consider a circuit containing an inductor (L), a capacitor (C), and a resistor (R) connected in series with a dc source of emf E. The instantaneous value of the current in the circuit is related to the circuit elements and the emf (E) of the battery through the equation

$$L\frac{dI}{dt} + RI + \frac{1}{C}\int I dt = E \qquad (3.42)$$

Applying LT, we get

$$L\left[sI(s) - I(0)\right] + RI(s) + \frac{1}{C}\left[\frac{I(s)}{s} + \frac{I(0)}{s} = \frac{E}{s}\right] \qquad (3.43)$$

We that for the given circuit, at t=0, the current is zero, i.e., I(0) =0. Eq(3.24) therefore becomes

$$\frac{E}{s} = L\left[sI(s)\right] + RI(s) + \frac{1}{C}\left[\frac{I(s)}{s}\right]$$

$$= \left[Ls + R + \frac{1}{Cs}\right]I(s)$$

$$\textit{or, } I(s) = \frac{E}{s\left[Ls + R + \frac{1}{Cs}\right]}$$

$$= \frac{E}{\left[Ls^2 + Rs + \frac{1}{C}\right]}$$

$$= \frac{E/L}{\left[s^2 + \frac{R}{L}s + \frac{1}{LC}\right]}$$

$$= \frac{E/L}{\left[s^2 + \frac{R}{L}s + \frac{1}{LC}\right]} \qquad (3.44)$$

$$= \frac{E/L}{\left[s^2 - (a+b)s + ab\right]}$$

$$\text{or, } I(s) = \frac{E/L}{\left[(s-a)(s-b)\right]} \qquad (3.45)$$

$$\text{where, } a + b = -\frac{R}{L},\ ab = \frac{1}{LC} \qquad (3.46)$$

$$\Rightarrow a = -\frac{R}{2L} + \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}$$

$$\Rightarrow b = -\frac{R}{2L} + \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}$$

$$\text{Again, } I(s) = \frac{E}{L}\frac{1}{a-b}\left[\frac{1}{s-a} - \frac{1}{s-b}\right] \qquad (3.47)$$

$$I(t) = L^{-1}\left[I(s)\right] \qquad (3.48)$$

$$= \frac{E}{L}\frac{1}{(a-b)}\int_0^\infty \left[\frac{1}{s-a} - \frac{1}{s-b}\right]e^{st}ds$$

$$= \frac{E}{L}\frac{1}{(a-b)}\left[e^{at}-e^{bt}\right]$$

$$= \frac{Ee^{-\frac{Rt}{2L}}}{L\sqrt{\frac{R^2}{L^2}-\frac{4}{LC}}}\left[e^{t\sqrt{\frac{R^2}{4L^2}-\frac{1}{LC}}}-e^{t\sqrt{\frac{R^2}{4L^2}-\frac{1}{LC}}}\right]$$

$$I(t) = \frac{2Ee^{-\frac{Rt}{2L}}}{\sqrt{R^2-\frac{4L}{C}}}\sinh\left(\sqrt{\frac{R^2}{4L^2}-\frac{1}{LC}}\right)t \qquad (3.49)$$

LR series circuit with constant voltage source

Let us consider a circuit containing an inductor (L) and a resistor (R) connected in series with a dc source of emf E. The insstantaneous value of the current in the circuit is related to the circuit elements and the emf E of the battery by the equation

$$L\frac{dI}{dt}+RI=E$$

Applying LT, we get

$$L\left[sI(s)-I(0)\right]+RI(s)=\frac{E}{s} \quad (3.51)$$

We that for the given circuit, at t = 0, the current is zero, i.e., I(0) =0. Eq(3.24) therefore becomes

$$\frac{E}{s} = L\left[sI(s)\right]+RI(s) \qquad (3.52)$$

$$= \left[Ls+R\right]I(s) \qquad (3.53)$$

$$\Rightarrow I(s) = \frac{E}{s(Ls+R)} \qquad (3.54)$$

$$= \frac{E}{s(s+R/L)} \qquad (3.55)$$

$$= \frac{E/L}{R/L}\left[\frac{1}{s} - \frac{1}{s + \frac{R}{L}}\right] \quad (3.56)$$

$$= \frac{E}{R}\left[\frac{1}{s} - \frac{1}{s + \frac{R}{L}}\right] \quad (3.57)$$

$$I(t) = L^{-1}\left[I(s)\right] \quad (3.58)$$

$$= \frac{E}{R}\int_0^{\infty}\left[\frac{1}{s} - \frac{1}{s + \frac{R}{L}}\right]e^{st}ds \quad (3.59)$$

$$I(t) = \frac{E}{R}\left[1 - e^{-\frac{Rt}{L}}\right] \quad (3.60)$$

Coupled differential equations of 1st order

Example: Solve the following system of coupled differential equations:

$$\frac{dx_1}{dt} = 3x_1 - 3x_2 + 2; x_1(0) = 1 \quad (3.61)$$

$$\frac{dx_1}{dt} = -6x_1 - t; x_2(0) = -1 \quad (3.62)$$

Takin Laplace transform of both differential equations, we have

$$sX(s) - x_1(0) = 3X_1(s) - 3X_2(s) + \frac{2}{s}$$

$$\Rightarrow \quad (s-3)X_1(s) + 3X_2(s) = 1 + \frac{2}{s} \quad (3.63)$$

$$sX_2(s) - x_2(0) = -6X_1(s) - \frac{1}{s^2}$$

$$\Rightarrow \quad 6X_1(s)+sX_2(s)=-1-\frac{1}{s^2} \tag{3.64}$$

Multiplying the Eq.(3.44) by s and Eq.(3.45) by -3 and adding we get

$$\left(s^2-3s-18\right)X_1(s) = 2+s+\frac{3s^2+3}{s^2}$$

$$\Rightarrow X_1(s) = \frac{s^3+5s^2+3}{s^2-3s-18}$$

$$= \frac{s^3+5s^2+3}{s^2(s+3)(s-6)}$$

$$X_1(s) = \frac{1}{108}\left(\frac{133}{s-6}-\frac{28}{s+3}+\frac{3}{s}-\frac{18}{s^2}\right) \tag{3.65}$$

Taking the inverse transform gives us the first solution,

$$x_1(t)=\frac{1}{108}\left(133e^{6t}-28e^{-3t}+3-18t\right) \tag{3.66}$$

Now to find the second solution we could go back up and eliminate X_1 to find the transform for X_2 and sometimes we would need to do that. However, in this case notice that the second differential equation is,

$$\frac{dx_2}{dt} = -6x_1-t$$

$$x_2(t) = \int\left[-6x_1(t)-t\right]dt$$

So, plugging the first solution in and integrating gives,

$$x_2(t) = \int\left[(-6)\left\{\frac{1}{108}\left(133e^{6t}-28e^{-3t}+3-18t\right)\right\}-t\right]dt$$

$$= -\frac{1}{108}\int\left(133e^{6t}-28e^{-3t}+3\right)dt$$

$$= -\frac{1}{108}\left(133e^{6t}-56e^{-3t}+18t\right)+c \tag{3.67}$$

Now, reapplying the second initial condition to get the constant of integration gives

$$-1 = -\frac{1}{108}(133+56)+c$$

$$\Rightarrow c = \frac{3}{4} \qquad (3.68)$$

The second solution is then,

$$x_2(t) = -\frac{1}{108}\left(133e^{6t}+56e^{-3t}+18t\right)+\frac{3}{4} \qquad (3.69)$$

Thus the solution of the given coupled equations are:

$$\left.\begin{aligned} x_1(t) &= \frac{1}{108}\left(133e^{6t}-28e^{-3t}+3-18t\right) \\ x_2(t) &= -\frac{1}{108}\left(133e^{6t}+56e^{-3t}+18t-81\right) \end{aligned}\right\} \qquad (3.70)$$

Exercises on application of Laplace transform

1. Find $f * g$ where $f(t) = e^{-t}$ and $g(t) = \sin t$.

   Hints: Convolution: $(f * g)(t) = \int_0^t (t-s)g(s)ds$

   Solution:

   $$(f*g)(t) = \int_0^t e^{-(t-s)}\sin s ds = \frac{1}{2}\left[e^{-(t-s)}(\sin s-\cos s)\right]_0^t$$

   $$= \frac{1}{2}\left[(\sin t-\cos t)+e^{-t}\right]$$

2. Express the solution to the initial value problem:

   $y'+\alpha y = g(t),\ y(0) = y_0,$ in terms of a convolution integral.

   Solution:

   Solving this initial value problem by the method of integrating factor we find

   $$y(t) = e^{-\alpha t}y_0+\int_0^t e^{-\alpha(t-s)}\sin s ds = e^{-\alpha t}y_0+e^{-\alpha t}*g(t)$$

3. For circuit containing a capacitor and a resistance (C-R) with source of constant EMF E and a key (K) all in series the the instantaneous charge on either plate of the capacitor is Q. Solve the corresponding dfferential equation

$$R\frac{dQ}{dt}+\frac{Q}{C}=E$$

to find $Q(t)$ using Laplace transform method.

Chapter-end exercise

1. Find the Laplace transform of $F(t)$ = exp(kt).

2. Show that $\mathcal{L}\{\sin \omega t\} = \frac{\omega}{s^2+\omega^2}$ $(s > 0)$.

3. Find the Laplace transform of the function

$$F(t)=\begin{cases} 0 & t\le 0 \\ \sin t & 0\le t\le 2\pi \\ \sin t+\cos t & t>2\pi \end{cases}$$

4. Use Laplace transforms to solve the initial value problem:

$y''+3y'+2y=e^{-t},\ y'(0)=0$

5. Use Laplace transforms to solve the initial value problem:

$y'+2y=26\sin 3t,\ y(0)=3.$

6. Use Laplace transforms to solve the initial value problem:

$y''+3y'+2y=6e^{-t},\ y(0)=1,\ y'(0)=2.$

7. Use Laplace transforms to solve the initial value problem:

$y''-2y'+y=e^{2t},\ y(0)=0,\ y'(0)=0$

8. Obtain general solution of heat flow along an infinite bar using Laplace transform.

Unit 4 ❑ Tensors

Structure

4.1 Objectives

4.2 Introduction

4.3 Tensors as multilinear transformations

4.4 Examples of tensors

4.5 Components of a tensor in basis

4.6 Symmetric and antisymmetric tensors

4.7 The completely symmetric and antisymmetric tensor

4.8 Summation convention

4.9 Inner product of vectors and the metric tensor

4.10 Coordinate systems and coordinate basis vectors

4.11 Reciprocal coordinate basis

4.12 Components of metric

4.13 Change of basis

4.14 Change of tensor components

4.15 Example : Intertial coordinates

4.16 Lorentz transformations as coordinate transformations

4.17 Elelctro-magnetic tensor and

4.1 Objectives

By the end of this unit, students will be able to:

- understand concept of tensor variables and difference from scalar or vector variables.
- understand the reason why the tensor analysis is used and explain usefulness of the tensor analysis.
- derive base vectors, metric tensors and strain tensors in an arbi-trary coordinate system.
- understand the meaning of symmetric, antisymmetric and completely antisymmetric tensors.

- understand summation convetion
- learn about change of tensor components under change of coor-dinate systems,
- understand inertial coordinates & bases in Minkowski space
- know Elelctro-magnetic tensor and change in its components un-der Lorentz transformations etc.

4.2 Introduction

Laws of Physics must be independent of any particular coordinate systems used in describing them mathematically if they are to be valid. A study of the consequences of this requirement leads to tensor analysis, of great use in general relativity theory, differential geometry, mechanics, elasticity, hydrodynamics, electromagnetic theory, and numerous other fields of science and engineering.

In three-dimensional space, a point is a set of three numbers, called coordinates, determined by specifying a particular coordinate system or frame of reference. For example, (x, y, z), (r, ϕ, z), (r, θ, ϕ) are coor-dinates of a point in rectangular, cylindrical, and spherical coordinate systems respectively. A point in N-dimensional space is, by analogy, a set of N numbers denoted by $\left(x^1, x^2, ..., x^N\right)$ where 1, 2, ..., N are taken not as exponents but as superscripts, a policy which will prove useful. The fact that we cannot visualize points in spaces of dimension higher than three has of course nothing whatsoever to do with their existence. Thus, in mathematics, a tensor is a geometric object that maps in a multi-linear manner geometric vectors, scalars, and other tensors to a resulting tensor. Vectors and scalars are the simplest tensors. Vectors from the dual space of the vector space, which supplies the geometric vectors, are also included as tensors.

4.3 Tensors as multilinear transformations (functionals) on vectors

Coordinate transformations

Let $\left(x^1, x^2, ..., x^N\right)$ and $\left(\bar{x}^1, \bar{x}^2, ..., \bar{x}^N\right)$ be coordinates of a point in two different frames of reference. Suppose there exists N independent relations between the coordinates of the two systems having the form

$$\left.\begin{aligned}
\bar{x}^1 &= \bar{x}^1\left(x^1, x^2, \ldots, x^N\right)\\
\bar{x}^2 &= \bar{x}^2\left(x^1, x^2, \ldots, x^N\right)\\
. &\quad . \ldots\ldots\ldots\ldots\\
. &\quad . \ldots\ldots\ldots\ldots\\
. &\quad . \ldots\ldots\ldots\ldots\\
\bar{x}^N &= \bar{x}^N\left(x^1, x^2, \ldots, x^N\right)
\end{aligned}\right\} \qquad (4.1)$$

which we can indicate briefly by

$$\bar{x}^k = \bar{x}^k\left(x^1, x^2, \ldots, x^N\right), k = 1, 2, 3, \ldots, N \qquad (4.2)$$

where it is supposed that the functions involved are single-valued, continuous, and have continuous derivatives. Then conversely to each set of coordinates $\left(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^N\right)$ there will correspond a unique set $\left(x^1, x^2, \ldots, x^N\right)$ given by

$$x^k = x^k\left(\bar{x}^1, \bar{x}^2, \ldots, \bar{x}^N\right), k = 1, 2, 3, \ldots, N \qquad (4.3)$$

Relations (4.2) or (4.3) define the transformation of coordinates from one frame of reference to another.

Contravariant and covariant vectors

If N quantities $A^1, A^2, \ldots, A^N$, in a coordinate system $(x^1, x^2, \ldots, x^N)$ are related to N other quantities $\bar{A}^1, \bar{A}^2, \ldots, \bar{A}^N$ in another coordinate system $(x^1, x^2, \ldots, x^N)$ by the transformation equations

$$\bar{A}^p = \sum_{q=1}^{N} \frac{\partial \bar{x}^p}{\partial x^q} A^q, \; p = 1, 2, 3, \ldots, N \qquad (4.4)$$

which by the adopted conventions can simply be written as

$$\bar{A}^p = \frac{\partial \bar{x}^p}{\partial x^q} A^q \qquad (4.5)$$

they are called components of a contravariant vector or contravariant tensor of the first rank or first order. If N quantities $A_1, A_2,...,A_N$ in a coordinate system $(x^1, x^2,..., x^N)$ are related to N other quantities $A_1, A_2,..., A_N$ in another coordinate system $\left(\overline{x}^1, \overline{x}^2,...\overline{x}^N\right)$ by the transformation equations

$$\overline{A}_p = \sum_{q=1}^{N} \frac{\partial x^q}{\partial \overline{x}^p} A_q, \;\; p = 1,2,3,...,N \tag{4.6}$$

or,

$$\overline{A}_p = \frac{\partial x^q}{\partial \overline{x}^p} A_q \tag{4.7}$$

they are called components of a covariant vector or covariant tensor of the first rank or first order. Note that a superscript is used to indicate contravariant components whereas a subscript is used to indicate covariant components; an exception occurs in the notation for coordinates. Instead of speaking of a tensor whose components are A^p or A_p we shall often refer simply to the tensor A^p or A_p. No confusion should arise from this.

Contravariant, Covariant and Mixed Tensors

If N^2 quantities A^{qs} in a coordinate system $(x^1, x^2,...,x^N)$ are related to N^2 other quantities $\overline{A}^{pr}$ in another coordinate system $\left(\overline{x}^1, \overline{x}^2,..., \overline{x}^N\right)$ by the transformation equations

$$\overline{A}^{pr} = \sum_{s=1}^{N}\sum_{q=1}^{N} \frac{\partial \overline{x}^p}{\partial x^q} \frac{\partial \overline{x}^r}{\partial x^s} A^{qs}; p,r = 1,2,3,...,N \tag{4.8}$$

or

$$\overline{A}^{pr} = \frac{\partial \overline{x}^p}{\partial x^q} \frac{\partial \overline{x}^r}{\partial x^s} A^{qs} \tag{4.9}$$

by the adopted conventions, they are called contravariant components of a tensor of the second rank or rank two.

The N^2 quantities A^{qs} are called covariant components of a tensor of the second rank if

$$\bar{A}_{pr} = \frac{\partial x^q}{\partial \bar{x}^p}\frac{\partial x^s}{\partial \bar{x}^r}A_{qs} \tag{4.10}$$

Similarly, the N^2 quantities A_s^q are called components of a mixed tensor of the second rank if

$$\bar{A}_r^p = \frac{\partial \bar{x}^p}{\partial x^q}\frac{\partial x^s}{\partial \bar{x}^r}A_s^q \tag{4.11}$$

Kronecker delta

Kronecker delta written δ_k^j, is defined by

$$\delta_k^j = \begin{cases} 1 \text{ if } j = k \\ 0 \text{ if } \neq k \end{cases} \tag{4.12}$$

As its notation indicates, it is a mixed tensor of the second rank.

Tensor of rank greater than two

Tensors of rank greater than two (i.e., rank > 2) are easily defined. For example, A_{lm}^{ijk} are the components of a mixed tensor of rank 5, contravariant of order 3, and covariant of order 2, if they transform according to the relations

$$A_{lm}^{ijk} = \frac{\partial \bar{x}^i}{\partial x^p}\frac{\partial \bar{x}^j}{\partial x^q}\frac{\partial \bar{x}^k}{\partial x^r}\frac{\partial x^s}{\partial \bar{x}^l}\frac{\partial x^t}{\partial \bar{x}^m}A_{st}^{pqr} \tag{4.13}$$

Scalar or Invariants

Suppose ψ is a function of the coordinates x^k, and let $\bar{\psi}$ denotes the functional value under a transformation to a new set of coordinates $\bar{x}^k$. Then ψ is called a scalar or invariant with respect to the coordinate transformation if $\psi = \bar{\psi}$. A scalar or invariant is also called a tensor of rank zero.

Tensor Fields

If to each point of a region in N-dimensional space there corresponds a definite tensor, we say that a tensor field has been defined. This is a vector field or a scalar field according to the tensor is of rank one or zero. It should be noted that a tensor or a tensor field is not just the set of its components in one special coordinate system but all the possible sets under any transformation of coordinates.

Fundamental operations with tensor

Addition

The sum of two or more tensors of the same rank and type (i.e. same number of contravariant indices and the same number of covariant indices) is also a tensor of the same rank and type. Thus if A_r^{pq} and B_r^{pq} are tensors, then $C_r^{pq} = A_r^{pq} + B_r^{pq}$ is also a tensor. The addition of tensors is commutative and associative.

Subtraction

The difference of two tensors of the same rank and type is also a tensor of the same rank and type. Thus if A_r^{pq} and B_r^{pq} are tensors, then $D_r^{pq} = A_r^{pq} - B_r^{pq}$ is also a tensor.

Outer Multiplication

The product of two tensors is a tensor whose rank is the sum of the ranks of the given tensors. This product which involves ordinary multiplication of the components of the tensor is called the outer product. For example, $A_r^{pq} B_m^l = C_{rm}^{pql}$ is the outer product of A_r^{pq} and B_m^l. However, note that not every tensor can be written as a product of two tensors of lower rank. For this reason, the division of tensors is not always possible.

Contraction

If one contravariant and one covariant index of a tensor are set equal, the result indicates that a summation over the equal indices is to be taken according to the summation convention. This resulting sum is a tensor of rank two less than that of the original tensor. The process is called contraction.

For example, in the tensor of rank 5, A_{lm}^{ijk}, set $m = k$ to obtain $A_{lk}^{ijk} = B_l^{ij}$, a tensor of rank 3. Further, by setting $l = j$ we obtain $B_j^{ij} = B^i$ a tensor of rank 1.

Inner Multiplication

By the process of outer multiplication of two tensors followed by a contraction, we obtain a new tensor called an inner product of the given tensors. The process is called inner multiplication. For example, given the tensors A_k^{ij} and B_{qr}^{p}, the outer product is $A_k^{ij}B_{qr}^{p}$. Letting $k = p$, we obtain the inner product $A_p^{ij}B_{qr}^{p}$. Letting $k = p$ and $q = j$, another inner product $A_p^{ij}B_{jr}^{p}$ is obtained. Inner and outer multiplication of tensors is commutative and associative.

Quotient Law

Suppose it is not known whether a quantity X is a tensor or not. If an inner product of X with an arbitrary tensor is itself a tensor, then X is also a tensor. This is called the quotient law.

Notes on Quotient Rule

In the tensor analysis, it is often necessary to ascertain whether a given quantity is tensor or not and if it is tensor we have to find its rank. The direct method requires us to find out if the given quantity obeys the transformation law or not. In practice this is troublesome and a similar test is provided by law is known as Quotient law. Generally, we can write,

$$KA = B \tag{4.14}$$

Here A and B are tensors of known rank and K is an unknown quantity. The Quotient Rule gives the rank of K. For example

$$I\vec{\omega} = \vec{L} \tag{4.15}$$

Here $\vec{L}$ and $\vec{\omega}$ are known vectors, then Quotient Rule shows that I is a second rank tensor. Similarly,

$$m\vec{a} = \vec{F} \tag{4.16}$$

$$\sigma\vec{E} = \vec{J} \tag{4.17}$$

$$\chi\vec{E} = \vec{P} \tag{4.18}$$

all establish the second rank tensor of m, σ, χ. The well known Quotient Rules are

$$K_i A_i = B \tag{4.19}$$

$$K_{ij} A_j = B_i \tag{4.20}$$

$$K_{ij}A_{jk} = B_{ik} \tag{4.21}$$

$$K_{ijkl}A_{ij} = B_{kl} \tag{4.22}$$

$$K_{ij}A_{k} = B_{ijk} \tag{4.23}$$

In each case, A and B are known tensors of rank indicated by the number of indices and A is arbitrary whereas in each case K is an unknown quantity. We have to establish the transformation properties of K.

The Quotient Rule asserts that if the equation of interest holds in all Cartesian co-ordinate systems, K is a tensor of indicated rank. The importance of physical theory is that the quotient rule establishes the tensor nature of quantities. There is an interesting idea that if we reconsider Newtons equations of motion

$$m\vec{a} = \vec{F}$$

based on the quotient rule that, if the mass is a scalar and the force a vector, then you can show that the acceleration × a is a vector. In other words, the vector character of the force as the driving term imposes its vector character on the acceleration, provided the scale factor m is scalar. This will first make us think that it contradicts the idea given in the introduction. But when we say m is scalar immediately we are considering that $\vec{a}$ and $\vec{F}$ have the same directions which make m scalar. Now let us prove each equation and find the nature of K.

Proof of the quotient rules

1. The quotient rule 4.19:

$$\mathbf{KA_i = B}$$

Proof: Taking prime on both sides

$$K'A_i' = B'$$

Here A has one index and hence it is a vector. Using the transformation equation for a vector

$$K'\left(\frac{\partial x_i'}{\partial x_j}\right)A_j = B$$

because $B' = B$, since it is a scalar. Now using the given rule RHS is modified and we get

$$\left.\begin{aligned} K'\left(\frac{\partial x_i'}{\partial x_j}\right)A_j &= KA_j \\ or, \left(K'\frac{\partial x_i'}{\partial x_j} - K\right)A_j &= 0 \end{aligned}\right\}$$

. Now Aj cannot be zero since it is a component and if it vanishes the law itself does not exist. Hence the quantity within the bracket vanishes.

$$K = K'\left(\frac{\partial x_i'}{\partial x_j}\right) \tag{4.24}$$

Here the transformation is with one coefficient and thus K is a first rank tensor.

2. The quotient rule 4.20:

$$\mathbf{KA_j = B_i}$$

Proof: Now we will proceed as in the case 1, taking prime on both sides

$$\left.\begin{aligned} K'A_j' &= B_i' \\ \Rightarrow K'\left(\frac{\partial x_j'}{\partial x_s}\right)A_s &= \left(\frac{\partial x_i'}{\partial x_p}\right)B_p \\ \Rightarrow K'\left(\frac{\partial x_j'}{\partial x_s}\right)A_s &= \left(\frac{\partial x_i'}{\partial x_p}\right)A_s \\ \Rightarrow \left(K'\frac{\partial x_j'}{\partial x_s} - \frac{\partial x_i'}{\partial x_p}K\right)A_s &= 0 \\ \Rightarrow K &= K'\left(\frac{\partial x_j'}{\partial x_s}\frac{\partial x_p}{\partial x_i}\right) \end{aligned}\right\} \tag{4.25}$$

Thus K is a second rank tensor.

3. The quotient rule 4.21:

$$\mathbf{KA}_{jk} = \mathbf{B}_{ik}$$

Proof:

Taking prime

$$\left.\begin{aligned} K'A'_{jk} &= B'_{ik} \\ \Rightarrow K'\left(\frac{\partial x'_j}{\partial x_p}\frac{\partial x'_k}{\partial x_q}\right)A_{pq} &= \left(\frac{\partial x'_i}{\partial x_r}\frac{\partial x'_k}{\partial x_q}\right)B_{rq} \\ \Rightarrow K'\left(\frac{\partial x'_j}{\partial x_p}\right)A_{pq} &= \left(\frac{\partial x'_i}{\partial x_r}\right)KA_{pq} \\ \Rightarrow\left(K'\frac{\partial x'_j}{\partial x_p}-\frac{\partial x'_i}{\partial x_r}K\right)A_{pq} &= 0 \\ \Rightarrow K &= K'\left(\frac{\partial x'_j}{\partial x_p}\frac{\partial x_r}{\partial x'_i}\right) \end{aligned}\right\} \quad (4.26)$$

K is a second rank tensor.

4. The quotient rule 4.22:

$$\mathbf{KA}_{ij} = \mathbf{B}_{kl}$$

Proof:

Taking prime

$$\left.\begin{aligned} K'A'_{jk} &= B'_{kl} \\ \Rightarrow K'\frac{\partial x'_j}{\partial x_p}\frac{\partial x'_j}{\partial x_q}A_{pq} &= \frac{\partial x'_k}{\partial x_r}\frac{\partial x'_l}{\partial x_s}B_{rs} \\ \Rightarrow K'\frac{\partial x'_j}{\partial x_p}\frac{\partial x'_j}{\partial x_q}A_{pq} &= \frac{\partial x'_k}{\partial x_r}\frac{\partial x'_l}{\partial x_s}K\,A_{pq} \\ \Rightarrow\left(K'\frac{\partial x'_j}{\partial x_p}\frac{\partial x'_j}{\partial x_q}-\frac{\partial x'_k}{\partial x_r}\frac{\partial x'_l}{\partial x_s}K\right)A_{pq} &= 0 \\ \Rightarrow K &= K'\left(\frac{\partial x'_i}{\partial x_p}\frac{\partial x'_j}{\partial x_q}\frac{\partial x_r}{\partial x'_k}\frac{\partial x_s}{\partial x_l}\right) \end{aligned}\right\} \quad (4.27)$$

K is a 4^{th} rank tensor.

5. The quotient rule 4.23:

$$\mathbf{KA_k = B_{ijk}}$$

Proof:

Taking prime

$$\left.\begin{aligned} K'A'_k &= B'_{ijk} \\ \Rightarrow K'\frac{\partial x'_k}{\partial x_r}A_r &= \frac{\partial x'_i}{\partial x_p}\frac{\partial x'_k}{\partial x_q}\frac{\partial x'_k}{\partial x_r}B_{pqr} \\ \Rightarrow K'A_r &= \frac{\partial x'_i}{\partial x_p}\frac{\partial x'_j}{\partial x_q}K\,A_r \\ \Rightarrow \left(K'\frac{\partial x'_i}{\partial x_p}\frac{\partial x'_j}{\partial x_q}K\right)A_r &= 0 \\ \Rightarrow K &= K'\left(\frac{\partial x_p}{\partial x'_i}\frac{\partial x_q}{\partial x_j}\right) \end{aligned}\right\} \tag{4.28}$$

Hence K is a second rank tensor.

Exercises

1. The double summation $K_{ij}A_i = B_j$ is invariant for any two vectors A_i and B_j. Prove that K_{ij} is a second-rank tensor.
2. The equation $K_{ij}A_{jk} = B_{ik}$ holds for all orientations of the coordinate system. If $\vec{A}$ and $\vec{B}$ are arbitrary second rank tensors show that K is a second rank tensor.

Matrices

A matrix of order m by n is an array of quantities a_{ij}, called elements, arranged in m rows and n columns enclosed between pair of parentheses or in brackets and generally denoted by

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \text{ or } \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$$

or in abbreviated form by (a_{ij}) or $[a_{ij}]$, $i = 1,...,m$; $j = 1,...,n$. If $m = n$ the matrix is a square matrix of order m by m or simply m; if $m = 1$ it is a row matrix or row vector; if $n = 1$ it is a column matrix or column vector. The diagonal of a square matrix containing the elements a_{11}, a_{22},...,a_{mm} is called the principal or main diagonal. A square matrix whose elements are equal to one in the principal diagonal and zero else is called a unit matrix and is denoted by I. A null matrix, denoted by O, is a matrix all of whose elements are zero.

4.4 Examples of tensors

Moment of Inertia Tensor

Finding the components of the moment of inertia is the simplest example given in many textbooks introducing nine component physical quantity. Consider a rigid body rotating with fixed angular velocity ω about an axis that passes through the origin (see Figure 4.1). Let $\mathbf{r}_i$ be the position vector of the ith mass element, whose mass is m_i. We expect this position vector to precess about the axis of rotation (which is parallel to ω) with angular velocity ω. We, therefore, have

$$\frac{d\mathbf{r}_i}{dt} = \omega \times \mathbf{r}_i$$

Thus, the above equation specifies the velocity, $\mathbf{v}_i = d\mathbf{r}_i/dt$, of each mass element as the body rotates with fixed angular velocity É about an axis passing through the origin. The total angular momentum of

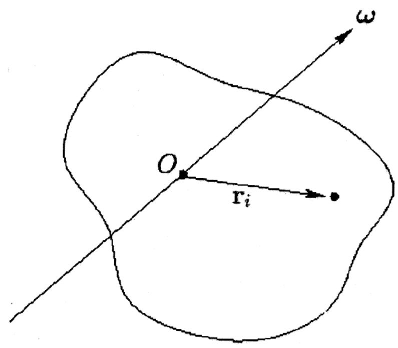


Figure 4.1: A rigid body

the body (about the origin) is written

$$L = \sum_{i=1,N} m_i \mathbf{r}_i \times \frac{d\mathbf{r}_i}{dt} = \sum_{i=1,N} m_i \mathbf{r}_i \times (w \times \mathbf{r}_i)$$
$$= \sum_{i=1,N} m_i \left[(\mathbf{r}_i^2 w - \mathbf{r}_i . w)\mathbf{r}_i \right], \tag{4.29}$$

where use has been made of Equation (4.14), and some standard vector identities. The above formula can be written as a matrix equation of the form

$$\begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix}, \tag{4.30}$$

where

$$I_{xx} = \sum_{i=1,N} \left(y_i^2 + z_i^2\right) m_i = \int (y^2 + z^2) dm, \tag{4.31}$$

$$I_{yy} = \sum_{i=1,N} \left(x_i^2 + z_i^2\right) m_i = \int (x^2 + z^2) dm, \tag{4.32}$$

$$I_{zz} = \sum_{i=1,N} \left(x_i^2 + y_i^2\right) m_i = \int (x^2 + y^2) dm, \tag{4.33}$$

$$I_{xy} = I_{yx} = -\sum_{i=1,N} x_i y_i m_i = -\int x\, y\, dm, \tag{4.34}$$

$$I_{yz} = I_{zy} = -\sum_{i=1,N} y_i z_i m_i = -\int y\, z\, dm, \tag{4.35}$$

$$I_{xz} = I_{zx} = -\sum_{i=1,N} x_i z_i m_i = -\int x\, z\, dm, \tag{4.36}$$

Here, I_{xx} is called the moment of inertia about the x-axis, I_{yy} the moment of inertia about the y-axis, I_{xy} the xy product of inertia, I_{yz} the yz product of inertia, etc. The matrix of the I_{ij} values is known as the moment of inertia tensor.

Note that each component of the moment of inertia tensor can be written as either a sum over separate mass elements, or as an integral over infinitesimal mass elements. In the integrals, $dm = \rho\, dV$, where ρ is the mass density, and dV a volume element. Equation (4.15) can be written more succinctly as

$$\mathbf{L} = \bar{\mathbf{I}}\omega \tag{4.37}$$

Here, it is understood that **L** and $\acute{E}$ are both column vectors, and $\bar{\mathbf{I}}$ is the matrix of the I_{ij} values. Note that $\bar{\mathbf{I}}$ is a real symmetric matrix: i.e., $I_{ij}^* = I_{ij}$ and $I_{ji} = I_{ij}$. In general, the angular momentum vector, **L**, obtained from Equation (4.22), points in a different direction to the angular velocity vector, ω. In other words, **L** is generally not parallel to ω.

Finally, although the above results were obtained assuming a fixed angular velocity, they remain valid at each instant in time if the angular velocity varies.

Excercises

1. Find the principal axes of rotation and the principal moments of inertia for a thin uniform rectangular plate of mass m and dimensions $2a$ by a for rotation about axes passing through (i) the center of mass, and (ii) a corner.

Dielectric susceptibility tensor

When an electric field passes through a dielectric medium, it causes polarization for the medium, and we define the electric susceptibility χ_e at some point in the dielectric as:

$$\mathbf{P} = \epsilon_0\, \chi_e \mathbf{E} \tag{4.38}$$

Where **P** is the electric dipole moment per unit volume and **E** is the total electric field at that point.

Well, if the dielectric is "isotropic", meaning **P** is independent of the orientation of the E-field, χ_e will be a scalar. However, if the dielectric is "an-isotropic", χ_e will be a rank-2 tensor and **P** and **E** will not necessarily be collinear.

The question is: Why shall **P** and **E** be non-collinear? How does it happen (the physical process)?

The expression of Px, for example, will be

$$P_x = \epsilon_0\, \chi_{xx} E_x + \epsilon_0\, \chi_{xy} E_y + \epsilon_0\, \chi_{xz} E_z \tag{4.39}$$

Thus the x-component of **P** depend on the y- and z-components of **E**. And the same is true for other components also. Hence,

$$\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix} = \epsilon_0 \begin{pmatrix} \chi_{xx} & \chi_{xy} & \chi_{xz} \\ \chi_{yx} & \chi_{yy} & \chi_{yz} \\ \chi_{zx} & \chi_{zy} & \chi_{zz} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}, \tag{4.40}$$

which can be expressed as

$$\mathbf{P} = \epsilon_0 \tilde{\chi}\, \mathbf{E}. \tag{4.41}$$

where $\tilde{\chi}$ is called the dielectric susceptibility tensor.

4.5 Components of a tensor in basis

Let us consider a second-order Cartesian tensor defined as

$$T = T_{ij}\mathbf{e_i} \otimes \mathbf{e_j} \tag{4.42}$$

The coefficients T_{ij} are the components of T. A tensor exists independent of any coordinate system. The tensor will have different components in different coordinate systems. The tensor T has components T_{ij} with respect to basis $\mathbf{e_i}$ and components T'_{ij} with respect to basis $\mathbf{e'_i}$, i.e.,

$$T = T_{pq}\mathbf{e_p} \otimes \mathbf{e_q} = T'_{ij}\mathbf{e'_i} \otimes \mathbf{e'_j} \tag{4.43}$$

Applying basic rules of coordinate transformation of tensors we have

$$T_{pq}\mathbf{e_p} \otimes \mathbf{e_q} = T_{pq}Q_{ip}Q_{jq}\mathbf{e'_i} \otimes \mathbf{e'_j} = T'_{ij}\mathbf{e'_i} \otimes \mathbf{e'_j}, \tag{4.44}$$

giving us

$$T'_{ij} = T_{pq}Q_{ip}Q_{jq}. \tag{4.45}$$

Similarly,

$$T'_{ij}\mathbf{e'_i} \otimes \mathbf{e'_j} = T'_{ij}Q_{ip}Q_{jq}\mathbf{e_p} \otimes \mathbf{e_q} = T_{pq}\mathbf{e_p} \otimes \mathbf{e_q} \tag{4.46}$$

gives the relation

$$T_{pq} = T'_{ij}Q_{ip}Q_{jq} = Q_{ip}Q_{jq}T'_{ij} \tag{4.47}$$

Equations (4.45) and (4.47) are the transformation rules for changing second-order tensor components under change of basis. In general Cartesian tensors of higher order can be expressed as

$$T = T_{ijk....n}\mathbf{e_i} \otimes \mathbf{e_j} \otimes \mathbf{e_j} \cdots \otimes \mathbf{e_n} \tag{4.48}$$

and the components transform according to

$$\left.\begin{aligned} T'_{ijk...} &= Q_{ip}Q_{jq}Q_{kr} \cdots T_{pqr...} \\ T_{pqr...} &= Q_{ip}Q_{jq}Q_{kr} \cdots T_{ijk...} \end{aligned}\right\} \tag{4.49}$$

4.6 Symmetric and antisymmetric tensors

Symmetric tensors

A tensor is called symmetric with respect to two contravariant or two covariant indices if its components remain unaltered upon interchange of the indices. Thus if $A^{ijk}_{lm} = A^{jik}_{lm}$ the tensor is symmetric in i and j. If a tensor is symmetric with respect to any two contravariant and any two covariant indices, it is called symmetric.

Antisymmetric (skew-symmetric) tensors

A tensor is called skew-symmetric with respect to two contravariant or two covariant indices if its components change sign upon interchange of the indices. Thus if $A^{ijk}_{lm} = A^{jik}_{lm}$ the tensor is skew-symmetric in i and j. If a tensor is skew-symmetric with respect to any two contravariant and any two covariant indices it is called skew-symmetric.

4.7 The completely symmetric and antisymmetric tensor

Completely symmetric tensors

Let's consider a tensor in d dimensions, meaning that each index runs from 1 to d. The rank (r) of the tensor is the number of indices that it has and the fact that it is totally symmetric means that $T_{...a...b...} = T_{...b...a...}$ for any pair of indices. The first example to look at is a tensor with two indices T_{ab}. This case is simple because it represents the components of a symmetric $d \times d$ matrix. The only independent components are the diagonal elements and the upper triangle because the lower triangle is determined from the upper one by the symmetry. There are of course d diagonal elements and we are left with $d^2 - d$ non-diagonal elements, which leads to $\frac{d(d-1)}{2}$ elements in the upper triangle. The total number of independent components is then

$$d + \frac{d(d-1)}{2} = \frac{d(d+1)}{2}. \tag{4.50}$$

The next example will be more complicated and will show the general idea. Consider a totally symmetric tensor of rank 3, T_{abc}. There are three types of components possible. We can first have things like T_{aaa} for a fixed number a from 1 to d and of course there are d such terms. We can next have terms like T_{aab} with $a \neq b$ and we don't care about T_{aba} or T_{baa} because they are related to T_{aab} by symmetry. The number of these components is $d(d-1)$ because you have d options for the first (repeated) index and then only $d-1$ possible choices for the remaining index. The last type of components is T_{abc} with $a \neq b \neq c$. There are $\binom{d}{3} = \frac{d(d-1)(d-2)}{3!}$ independent components of that form because we only need to choose three different 3! numbers between 1 and d and the order doesn't matter. The total number of independent components is then

$$d + d(d-1) + \frac{d(d-1)(d-2)}{3!} = \frac{d(d+1)(d+2)}{3!} \tag{4.51}$$

In general, the number of independent components of a totally symmetric tensor of rank r in d dimensions is

$$\binom{d+r-1}{r} = \frac{(d+r-1)!}{r!(d-1)!} = \frac{(d+r-1)(d+r-2)......(d+1)d}{r!} \tag{4.52}$$

Completely anti-symmetric tensors

It's possible to do the same kind of thing for totally anti-symmetric tensors that satisfy $T_{...a...b...} = -T_{...b...a...}$ for every pair of indices, but the analysis is easier. From the anti-symmetry we can already deduce that the value of all the indices for a non-zero component must be different because otherwise we would have $T_{...a...a...} = -T_{...a...a...}$ $\Rightarrow T_{...a...a...} = 0$. This then means that we can't have a non-trivial totally anti-symmetric tensor with $r > d$. For a generic $r \leq d$, since we can relate all the components that have the same set of values for the indices together by using the anti-symmetry, we only care about which numbers appear in the component and not the order. The number of independent components is then simply the number of ways of picking r numbers out of d without a specific order, which is

$$\binom{d}{r} = \frac{(d)!}{r!(d-r)!} \tag{4.53}$$

In particular, there is only one free component for an anti-symmetric tensor of rank d in d dimensions. Asking for $T_{12...d} = 1$ defines the well-known Levi-Civita tensor $\in_{a_1 a_2 ... a_d}$.

4.8 Summation convention

In writing an expression such as $a_1x^1 + a_2x^2 + ... + a_Nx^N$ we can use the short notation $\sum_{l=1}^{N} a_l x^l$. An even shorter notation is simply to write it as a_jx^j, where we adopt the convention that whenever an index (subscript or superscript) is repeated in a given term we are to sum over that index from 1 to N unless otherwise specified. This is called the summation convention. Clearly, instead of using the index l we could have used another letter, say p, and the sum could be written a_px^p. Any index which is repeated in a given term, so that the summation convention applies, is called a dummy index or umbral index. An index occurring only once in a given term is called a free index and can stand for any of the numbers 1, 2, ..., N such as k in equation (4.2) or (4.3), each of which represents N equations.

4.9 Inner product of vectors and the metric tensor

Inner product of vectors

An inner product is a generalization of the dot product. In a vector space, it is a way to multiply vectors together, with the result of this multiplication being a scalar. More precisely, for a real vector space, an inner product $<\cdot,\cdot>$ satisfies the following four properties. Let u, v, and w be vectors and α be a scalar, then:

1. $< u + v, w > = < u, w > + < v, w >$.

2. $< \alpha v,w >= \alpha < v,w >$.

3. $< v,w >=< w, v >$.

4. $< v, v > \geq 0$ and equal if and only if v=0.

The fourth condition in the list above is known as the positive-definite condition. Related thereto, note that some authors define an inner product to be a function $<\cdot,\cdot>$ satisfying only the first three of the above conditions with the added (weaker) condition of being (weakly) non-degenerate (i.e., if $<v,w>=0$ for all w, then $v \equiv 0$). In such literature, functions satisfying all four such conditions are typically referred to as positive-definite inner products (Ratcliffe 2006), though inner products which fail to be positive-definite are sometimes called indefinite to avoid confusion. This difference, though subtle, introduces several noteworthy phenomena: For example, inner products which fail to be positive-definite may give rise to "norms" which yield an imaginary magnitude for certain vectors (such vectors are called spacelike) and which induce "metrics" which fail to be actual metrics. The Lorentzian inner product is an example of an indefinite inner product.

A vector space together with an inner product on it is called an inner product space. This definition also applies to an abstract vector space over any field. Examples of inner product spaces include:

1. The real numbers R, where the inner product is given by $<x,y>=xy$.

2. The Euclidean space Rn, where the inner product is given by the dot product

$$< (x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n) \geq x_1y_1 + x_2y_2 + ...x_ny_n. \qquad (4.54)$$

3. The vector space of real functions whose domain is an closed interval [a,b] with inner product

$$\langle f, g >= \int_a^b f\,g dx. \qquad (4.55)$$

4. When given a complex vector space, the third property above is usually replaced by

$$< v,w >= \overline{< w,v >}, \qquad (4.56)$$

where $\bar{z}$ refers to complex conjugation. With this property, the inner product is called a Hermitian inner product and a complex vector space with a Hermitian inner product is called a Hermitian inner product space.

Every inner product space is a metric space. The metric is given by

$$g(v,w) = < v-w, v-w > . \tag{4.57}$$

If this process results in a complete metric space, it is called a Hilbert space. What's more, every inner product naturally induces a norm of the form $|x| = \sqrt{< x, x >}$, whereby it follows that every inner product space is also naturally a normed space. As noted above, inner products which fail to be positive-definite yield "metrics" - and hence, "norms" - are something different due to the possibility of failing their respective positivity conditions.

For example, n-dimensional Lorentzian Space (i.e., the inner product space consisting of R^n with the Lorentzian inner product) comes equipped with a metric tensor of the form

$$(ds)^2 = -dx_0^2 + dx_1^2 + \ldots + dx(n-1)^2 \tag{4.58}$$

and a squared norm of the form

$$|v|^2 = -v_0^2 + v_1^2 + \ldots + v(n-1)^2 \tag{4.59}$$

for all vectors $v = (v_0, v_1, \ldots, v(n-1))$. In particular, one can have negative infinitesimal distances and squared norms, as well as nonzero vectors whose vector norm is always zero. As such, the metric (respectively, the norm) fails to be a metric (respectively, a norm), thoughthey usually are still called such when no confusion may arise.

Metric tensor

An expression that represents the distance between two adjacent points is called a metric or line element. In three dimensional space the line element i.e. the distance between two adjacent points (x, y, z) and $(x + dx, y + dy, z + dz)$ in cartesian coordinate is given by $ds^2 = dx^2 + dy^2 + dz^2$. In terms of general curvilinear coordinates, the line element becomes

$$ds^2 = \sum_{\mu=1}^{3}\sum_{\nu=1}^{3} g_{\mu\nu} dq_\mu dx_\mu dq_\nu = g_{\mu\nu} dq_\mu dx_\mu dq_\nu \tag{4.60}$$

(using summation convention). This idea was generalized by Riemann to n-dimensional space. The distance between two neighbouring points with coordinate x^μ and $x^\mu + dx^\mu$ is given by

$$ds^2 = \sum_{\mu=1}^{n}\sum_{\nu=1}^{n} g_{\mu\nu} dx^{\mu} dx^{\nu} \qquad (4.61)$$

where the coefficients $g_{\mu\nu}$ are the functions of the coordinates x^{μ}, subject to the restriction g = determinant of $g_{\mu\nu} = |g_{\mu\nu}| \neq 0$. The quadratic differential form $g_{\mu\nu} dx^{\mu} dx^{\nu}$ is independent of the coordinate system and is called the Riemannian metric for n-dimensional space. The space which is characterised by Riemannian metric is called Riemannian space. Here the quantities $g_{\mu\nu}$ are components of a covariant symmetric tensor of rank two, called the metric tensor or fundamental tensor.

Exercises

1. If $dS^2 = g_{ij} dx^i dx^j$ is invariant, show that g_{ij} is a symmetric covariant tensor of rank 2.

Solutions

1. We have, $dS^2 = g_{ij} dx^i dx^j$.

   Since it is invariant, $dS^2 = \bar{g}_{ij} d\bar{x}^i d\bar{x}^j$.

   So, $\bar{g}_{ij} d\bar{x}^i d\bar{x}^j = \bar{g}_{kl} d\bar{x}^k d\bar{x}^l$.

Now applying inverse transformation law of dx^l and dx^m, we get,

$$\bar{g}_{ij} d\bar{x}^i d\bar{x}^j = g_{kl} \frac{\partial x^k}{\partial \bar{x}^i} d\bar{x} \frac{\partial x^l}{\partial \bar{x}^j} d\bar{x}^j = g_{kl} \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial x^l}{\partial \bar{x}^j} d\bar{x}^i d\bar{x}^j,$$

$$\Rightarrow \left(\bar{g}_{ij} - g_{kl} \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial x^l}{\partial \bar{x}^j} \right) d\bar{x}^i d\bar{x}^j = 0,$$

$\Rightarrow \bar{g}_{ij} = g_{kl} \dfrac{\partial x^k}{\partial \bar{x}^i} \dfrac{\partial x^l}{\partial \bar{x}^j} \Rightarrow$ the required transformation law for the second order covariant tensor. Hence, g_{ij} is a covariant tensor of rank two. Now, g_{ij} can be expressed in terms of symmetric and antisymmeric combinations as follows

$g_{ij} = \frac{1}{2}(g_{ij} + g_{ji}) + \frac{1}{2}(g_{ij} - g_{ji}) = A_{ij} + B_{ji}$ where $A_{ji} = \frac{1}{2}(g_{ij} + g_{ji})$ is a symmetric tensor and $B_{ij} = \frac{1}{2}(g_{ij} - g_{ji})$ is an antisymmetric tensor. Then $dS^2 = g_{ij} dx^i dx^j =$ $(A_{ij} + B_{ji}) dx^i dx^j$.

We can write, $B_{ij}dx^idx^j = B_{ij}dx^idx^j = B_{ji}dx^jdx^i = -B_{ij}dx^idx^j$ (interchanging dummy indices i and j and using the fact that B_{ij} is antisymmetric i.e., $B_{ij} = -B_{ji}$)

$2B_{ij}dx^idx^j = 0, \Rightarrow B_{ij} = 0$. i.e., $\frac{1}{2}(g_{ij} - g_{ji}) = 0, \, or \, g_{ij} = g_{ji}$, which shows that g_{ij} is symmetric.

4.10 Coordinate systems and coordinate basis vectors

Objectives

This section aims to introduce basis vectors, coordinate system, and representation of a vector in terms of basis vectors in an N-dimensional space V.

Outcomes

In this section learners will

- Learn to view a basis as a coordinate system on a subspace.
- Learn to compute the B-coordinates of a vector, compute the usual coordinates of a vector from its B-coordinates.
- Learn to find B-coordinates of a vector using its location on a nonstandard coordinate grid.

In this section, we interpret a basis $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of a subspace V as a coordinate system on V, and we learn how to write a vector in V in that coordinate system.

Writing a vector using B-coordinates

If $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, ..., \mathbf{v}_N\}$ be a basis of a subspace V, then $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + ... + c_N\mathbf{v}_N$ be a vector in V. The coefficients $c_1, c_2, c_3, ..., c_N$ are the coordinates of $\mathbf{v}$ with respect to the basis B.

Theorem

If $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for a subspace V, then any vector $\mathbf{A}$ in V can be written as a linear combination

$$\mathbf{A} = A_1\mathbf{v}_1 + A_2\mathbf{v}_2 + A_3\mathbf{v}_3 + ... + A_N\mathbf{v}_N \tag{4.62}$$

exactly in the same way.

Proof:

If B represents a basis for the vector space V, then B spans V and B is linearly independent. Since B spans V, we can write any $\mathbf{v}$ $(\equiv \vec{v})$ in V as a linear combination of $v_1, v_2, v_3, ..., v_N$. For uniqueness, suppose that we had two such expressions:

$$\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + c_3\mathbf{v_3} + ... + c_N\mathbf{v_N} \tag{4.63}$$

$$\mathbf{v} = c'_1\mathbf{v_1} + c'_2\mathbf{v_2} + c'_3\mathbf{v_3} + ... + c'_N\mathbf{v_N} \tag{4.64}$$

Subtracting the first equation from the second yields

$$\mathbf{v} - \mathbf{v} = (c_1 - c'_1)\mathbf{v_1} + (c_2 - c'_2)\mathbf{v_2} + (c_3 - c'_3)\mathbf{v_3} + ...+ (c_N - c'_N)\mathbf{v_N} = 0 \tag{4.65}$$

Since B is linearly independent, the only solution to the above equation is the trivial solution: all the coefficients must be zero. It follows that $(c_i - c'_i) = 0$ for all i, which proves that $c_1 = c'_1, c_2 = c'_2, ..., c_N = c'_N$.

Standard basis of R^N:

The standard coordinate vectors

$$\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}, \mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}, \mathbf{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ . \\ . \\ 0 \end{pmatrix}, ..., \mathbf{e_N} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ 1 \end{pmatrix} \tag{4.66}$$

form a basis for R^N. This is sometimes known as the standard basis. In particular, R^N has dimension N.

Standard basis of $R3$

A basis in a three-dimesional space, is any set of three linearly independent vectors $\mathbf{e_1}$, $\mathbf{e_2}$, $\mathbf{e_3}$ in that space. However, it is to be kept in mind that two linearly dependent vectors are collinear. For a given basis $\mathbf{e_1}$, $\mathbf{e_2}$, $\mathbf{e_3}$, every vector $\mathbf{v}$ in 3-D space has a unique representation of the form

$$\mathbf{v} = x\mathbf{e_1} + y\mathbf{e_2} + z\mathbf{e_3} \tag{4.67}$$

which can also be represented 3-D space as

$$\mathbf{v} = x\begin{pmatrix}1\\0\\0\end{pmatrix} + y\begin{pmatrix}0\\1\\0\end{pmatrix} + z\begin{pmatrix}0\\0\\1\end{pmatrix} \tag{4.68}$$

In the above representation, if × $e1$, × $e2$, × $e3$ are all orthogonal, of unit magnitude and constant, the coordinate system is called cartesian coordinate system.

Example:

$$\mathbf{v} = 2\begin{pmatrix}1\\0\\0\end{pmatrix} - 3\begin{pmatrix}0\\1\\0\end{pmatrix} + 4\begin{pmatrix}0\\0\\1\end{pmatrix} = 2\mathbf{e_1} - 3\mathbf{e_2} + 4\mathbf{e_3} \tag{4.69}$$

Eqs.(4.67) to (4.69) represent the relation between vector and bases in 3-dimensional space in cartesian coordinates.

Note: The N vectors $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, ..., $\mathbf{A}_N$ are called linearly independent if and only if

$$c_1\mathbf{A_1} + c_2\mathbf{A_2} + c_3\mathbf{A_3} + \ldots\ldots + c_N\mathbf{A_N} = 0. \tag{4.70}$$

Implies that each of the coefficients c_1, c_2, ..., c_N vanishes.

2-D Coordinate transforms of vectors

A vector cannot be described without a coordinate system. Let us have a vector **v** in the 2-D plane having components v_x, v_y along the x-axis and y-axis resapectively. Now we introduce a rotated coordinate system represented by the broken line as shown below, using x' and y'. The new system is rotated counter-clockwise by an angle, θ, from the initial coordinate system. It is to be noted that the vector itself remains unchanged, although it is described by different numerical values in the new coordinate system. In this case, the vector makes smaller angle with the x'-axis than with the y'-axis , so the i' component will be greater than the j' component. The vector in both the systems are represented as

$$\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j} = v_{x'}\mathbf{i'} + v_{y'}\mathbf{j'} \tag{4.71}$$

The 2-D vector transformation equations are

$$\left.\begin{aligned} v'_x &= v_x \cos\theta + v_y \sin\theta \\ v'_y &= -v_x \sin\theta + v_y \cos\theta \end{aligned}\right\} \tag{4.72}$$

This can be seen by noting that

- the part of v_x that lies along x' is $v_x \cos\theta$
- the part of v_y that lies along x' is $v_y \sin\theta$
- the part of v_x that lies along y' is $-v_x \sin\theta$
- the part of v_y that lies along y' is $v_y \cos\theta$

These four factors make up the four terms in the transformation equations. They are easily checked by setting q = 0° and q = 90°

- When $\theta = 0°$:
  $v'_x = v_x \cos\theta + v_y \sin\theta = v_x \cos 0° + v_y \sin 0° = v_x$
  $v'_y = -v_x \sin\theta + v_y \cos\theta = -v_x \sin 0° + v_y \cos 0° = v_y$
- When $\theta = 90°$:
  $v'_x = v_x \cos\theta + v_y \sin\theta = v_x \cos 90° + v_y \sin 90° = v_y$
  $v'_y = -v_x \sin\theta + v_y \cos\theta = -v_x \sin 90° + v_y \cos 90° = -v_x$
- When $\theta = 45°$:

  $$v'_x = v_x \cos\theta + v_y \sin\theta = v_x \cos 45° + v_y \sin 45° = \frac{1}{\sqrt{2}}\left[v_x + v_y\right]$$

  $$v'_y = -v_x \sin\theta + v_y \cos\theta = -v_x \sin 45° + v_y \cos 45° = \frac{1}{\sqrt{2}}[-v_x + v_y]$$

  It is obvious that, $v'_x > v'_y$ for $0° < \theta < 90°$.

Transformation matrix

It is more convenient to write (and work with) transformation equations using matrices.

$$\begin{bmatrix} v'_x \\ v'_y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} \tag{4.73}$$

The cos θ terms are on the matrix diagonal while the sin θ terms are offdiagonal. The above equation (4.73) can be written in matrix notation as

$$\mathbf{v}' = \mathbf{T}.\mathbf{v} \tag{4.74}$$

where **T** is chosen to represent the transformation matrix[1], i.e.,

$$\mathbf{T} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{4.75}$$

The rotation matrix R which arises when the coordinate axes remains fixed but the object rotates thtough angle θ:

$$\mathbf{R} = \mathbf{T}' \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}^{t} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{4.76}$$

1 Readers may get a little bit confused when one talks about the Transformation matrix and Rotation matrix. This confusion arises for not clarifying what is fixed and what is rotating. In the current discussion, it is the coordinate system that is rotating while the object remains fixed. So the term transformation matrix is used here to emphasize this.

However, in the situations in which the object rotates while the coordinate system remains fixed, the term rotation matrix will be used to emphasize that the object is rotating.

Another reason that causes this confusion is the amazing fact that each matrix (transformation and rotation) is just the transpose of the other! So they look extremely similar. In 2-D problems, the only practical difference is whether the minus sign in front of sin , is on the T_{12} term, or the T_{21} term.

4.11 Reciprocal coordinate basis

Let us consider three linearly independent vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ which are neither orthogonal nor have unit length (i.e they are non-orthonormal). Let any arbitrary vector **A** expanded w.r.t. this basis vectors has expansion coefficients components can be A^1, A^2, A^3, then

$$\mathbf{A} = A^1\mathbf{e}_1 + A^2\mathbf{e}_2 + A^3\mathbf{e}_3 \tag{4.77}$$

Now we shall aim towards finding the expansion coefficients. Concerning an orthonormal basis, the solution could be obtained by taking the dot product of the vector with each of the orthonormal basis vectors, i.e.,

$$A^i = \mathbf{A}.\mathbf{e}_i \tag{4.78}$$

When the basis is not orthogonal, the resolution of **A** will become less obvious. Considerable simplification to the problem can be achieved by introducing a new basis vectors called the **reciprocal basis** vectors. To different sets of basis vectors, say, ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$) and ($\mathbf{e}^1$, $\mathbf{e}^2$, $\mathbf{e}^3$) are said to be reciprocal basis if they satisfy

$$\mathbf{e}_i.\mathbf{e}^k = \delta_i^k. \tag{4.79}$$

In order to construct reciprocal basis from the ordinary one, it is to be noted that $\mathbf{e}^1$ must be perpendicular to both $\mathbf{e}_2$ and $\mathbf{e}_3$. Hence, we set

$$\mathbf{e}^1 = p(\mathbf{e}_2 \times \mathbf{e}_3) \tag{4.80}$$

The requirement that

$$\mathbf{e}_1 \cdot \mathbf{e}^1 = 1 \tag{4.81}$$

implies that

$$p\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3) = 1 \tag{4.82}$$

$$or,\ p = \frac{1}{\mathbf{e}_1.(\mathbf{e}_2 \times \mathbf{e}_3)} \tag{4.83}$$

Hence,

$$\mathbf{e}^1 = \frac{\mathbf{e}_2 \times \mathbf{e}_3}{\mathbf{e}_1.(\mathbf{e}_2 \times \mathbf{e}_3)} \tag{4.84}$$

In the similar way, we get

$$\mathbf{e}^2 = \frac{\mathbf{e}_3 \times \mathbf{e}_1}{\mathbf{e}_1.(\mathbf{e}_2 \times \mathbf{e}_3)} \tag{4.85}$$

$$\mathbf{e}^3 = \frac{\mathbf{e}_2 \times \mathbf{e}_1}{\mathbf{e}_1.(\mathbf{e}_2 \times \mathbf{e}_3)} \tag{4.86}$$

It is obvious that "an orthonormal basis is its own reciprocal basis", i.e.,

$$\left.\begin{array}{lcl} \hat{\mathbf{e}}^1 & = & \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}^2 & = & \hat{\mathbf{e}}_2 \\ \hat{\mathbf{e}}^3 & = & \hat{\mathbf{e}}_3 \end{array}\right\} \tag{4.87}$$

when $\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = \delta_i^j$ and $|\hat{\mathbf{e}}_i| = 1$. Taking the dot product of **A** with $\mathbf{e}^i$, we find

$$\mathbf{A}^i \cdot \mathbf{e}^i = A^i. \tag{4.88}$$

Hence in general we have

$$\mathbf{A} = (\mathbf{A}.\mathbf{e}^1)\mathbf{e}_1 + (\mathbf{A}.\mathbf{e}^2)\mathbf{e}_2 + (\mathbf{A}.\mathbf{e}^3)\mathbf{e}_3. \tag{4.89}$$

4.12 Components of metric in a coordinate basis and association with infinitesimal distance

In the rectangular cartesian coordinate basis (**i**, **j**, **k**), an infinitesimal distance between points located at $P(x, y, z)$ and $Q(x + dx, y + dy, z + dz)$ is expressed as

$$ds^2 = \mathbf{dr}.\mathbf{dr} = (\mathbf{i}dx + \mathbf{j}dy + \mathbf{k}dz).(\mathbf{i}dx + \mathbf{j}dy + \mathbf{k}dz) = dx^2 + dy^2 + dz^2. \tag{4.90}$$

In general, the distance between two neighbouring points with coordinate x^μ and $x^\mu + dx^\mu$ is given by

$$ds^2 \sum_{\mu=1}^{n} \sum_{\nu=1}^{n} g_{\mu\nu} dx^\mu dx^\nu = g_{\mu\nu} dx^\mu dx^\nu, \; [\mu, \nu = 1, 2, 3] \tag{4.91}$$

where the coefficients $g_{\mu\nu}$ are components of the covaiant metric tensor of rank two. Comparing (4.90) and (4.91) we have, $g_{11} = g_{22} = g_{33} = 1$, and $g_{\mu\nu} = 0$ for $\mu \neq \nu$. As $g_{\mu\nu}$ is a coavariant tensor of rank two, it transforms according to tensor transfomtion law as

$$\bar{g}_{\mu\nu} = \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} g_{\alpha\beta}. \tag{4.92}$$

When, $ds^2 = dx^2 + dy^2 + dz^2$,

$$g_{\mu\nu} = \sum_{\alpha=1}^{3} \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\alpha}{\partial \bar{x}^\nu} g_{\alpha\alpha}; \; [\text{since, } g_{\mu\nu} = 0 \text{ for } \mu \neq \nu]$$

$$or, \; \bar{g}_{\mu\nu} = \sum_{\alpha=1}^{3} \frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\alpha}{\partial \bar{x}^\nu}; \; [\text{as, } g_{\alpha\alpha} = 1] \tag{4.93}$$

Components of metric in spherical coordinates (r, θ, ϕ)

The transformation equatiion from cartesian to spherical polar coordinates are

$$\left.\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned}\right\} \qquad (4.94)$$

Let

$$\left.\begin{aligned} x^1 = x,\ x^2 = y\ x^3 = z \\ \bar{x}^1 = r,\ \bar{x}^2 = 0\ \bar{x}^3 = \phi \end{aligned}\right\} \qquad (4.95)$$

From Eq. (4.93), we have:

$$\begin{aligned} g_{11} &= \frac{\partial x^1}{\partial \bar{x}^1}\frac{\partial x^1}{\partial \bar{x}^1} + \frac{\partial x^2}{\partial \bar{x}^1}\frac{\partial x^2}{\partial \bar{x}^1} + \frac{\partial x^3}{\partial \bar{x}^1}\frac{\partial x^3}{\partial \bar{x}^1} \\ &= \left(\frac{\partial x}{\partial r}\right)^2 + \left(\frac{\partial y}{\partial r}\right)^2 + \left(\frac{\partial z}{\partial r}\right)^2, \text{ [using (4.82)]} \\ &= (\sin\theta\cos\phi)^2 + (\sin\theta\sin\phi)^2 + (\cos\theta)^2 \\ &= 1 \end{aligned}$$

Again from Eq. (4.93):

$$\begin{aligned} g_{22} &= \frac{\partial x^1}{\partial \bar{x}^2}\frac{\partial x^1}{\partial \bar{x}^2} + \frac{\partial x^2}{\partial \bar{x}^2}\frac{\partial x^2}{\partial \bar{x}^2} + \frac{\partial x^3}{\partial \bar{x}^2}\frac{\partial x^3}{\partial \bar{x}^2} \\ &= \left(\frac{\partial x}{\partial \theta}\right)^2 + \left(\frac{\partial y}{\partial \theta}\right)^2 + \left(\frac{\partial z}{\partial \theta}\right)^2, \text{ [using (4.82)]} \\ &= (r\cos\theta\cos\phi)^2 + (r\cos\theta\sin\phi)^2 + (-r\sin\theta)^2 \\ &= r^2 \end{aligned}$$

Again from Eq. (4.93):

$$
\begin{aligned}
g_{33} &= \frac{\partial x^1}{\partial \bar{x}^3}\frac{\partial x^1}{\partial \bar{x}^3} + \frac{\partial x^2}{\partial \bar{x}^3}\frac{\partial x^2}{\partial \bar{x}^3} + \frac{\partial x^3}{\partial \bar{x}^3}\frac{\partial x^3}{\partial \bar{x}^3} \\
&= \left(\frac{\partial x}{\partial \phi}\right)^2 + \left(\frac{\partial y}{\partial \phi}\right)^2 + \left(\frac{\partial z}{\partial \phi}\right)^2, \text{ [using (4.95)]} \\
&= (-r\sin\theta\sin\phi)^2 + (r\sin\theta\cos\phi)^2 + (0)^2 \\
&= r^2\sin^2\theta
\end{aligned}
$$

Again from Eq. (4.93):

$$
\begin{aligned}
g_{12} &= \frac{\partial x^1}{\partial \bar{x}^1}\frac{\partial x^1}{\partial \bar{x}^1} + \frac{\partial x^2}{\partial \bar{x}^1}\frac{\partial x^2}{\partial \bar{x}^2} + \frac{\partial x^3}{\partial \bar{x}^1}\frac{\partial x^3}{\partial \bar{x}^2} \\
&= \left(\frac{\partial x}{\partial r}\right)\left(\frac{\partial x}{\partial \theta}\right) + \left(\frac{\partial y}{\partial r}\right)\left(\frac{\partial y}{\partial \theta}\right) + \left(\frac{\partial z}{\partial r}\right)\left(\frac{\partial z}{\partial \theta}\right), \text{ [using (4.97)]} \\
&= (\cos\theta)(-\sin\theta) + (\sin\theta)(r\cos\theta) + (0)(0) \\
&= r\sin\theta\cos\theta + r\sin\theta\cos\theta \\
&= 0, [= g_{21}]
\end{aligned}
$$

Similarly, $g_{13} = g_{31} = 0$; $g_{23} = g_{32} = 0$. In general $g_{\mu\nu} = 0, for \mu \neq \nu$. Thus

$$
\begin{aligned}
ds^2 &= g_{\bar{\mu\nu}} d\bar{x}^\mu d\bar{x}^\nu \\
&= \sum_{\mu=1}^{3} g_{\bar{\mu\nu}} d\bar{x}^\mu d\bar{x}^\mu, \text{ [since, } g_{\mu\nu} = 0, \text{ for } \mu \neq \nu] \\
&= g_{\bar{11}}(d\bar{x}^1)^2 + g_{\bar{22}}(d\bar{x}^2)^2 + g_{\bar{33}}(d\bar{x}^3)^2 \\
&= (1)(dr)^r + (r^2)(d\theta)^2 + (1)(dz)^2 \\
&= dr^2 + r^2 d\theta^2 + dz^2
\end{aligned}
$$

Components of metric cylindrical coordinates (r,θ,ϕ)

The transformation equatiion from cartesian to spherical polar coordinates are

$$\left.\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}\right\} \tag{4.96}$$

Let

$$\left.\begin{aligned} x^1 &= x, \quad x^2 = y \quad x^3 = z \\ \bar{x}^1 &= r, \quad \bar{x}^2 = \theta \quad \bar{x}^3 = z \end{aligned}\right\} \tag{4.97}$$

From Eq. (4.93), we have:

$$\begin{aligned} g_{11} &= \frac{\partial x^1}{\partial \bar{x}^1}\frac{\partial x^1}{\partial \bar{x}^1} + \frac{\partial x^2}{\partial \bar{x}^1}\frac{\partial x^2}{\partial \bar{x}^1} + \frac{\partial x^3}{\partial \bar{x}^1}\frac{\partial x^3}{\partial \bar{x}^1} \\ &= \left(\frac{\partial x}{\partial r}\right)^2 + \left(\frac{\partial y}{\partial r}\right)^2 + \left(\frac{\partial z}{\partial r}\right)^2, \text{ [using (4.97)]} \\ &= (\cos\theta)^2 + (\sin\theta)^2 + (0)^2, \text{ [using (4.96)]} \\ &= 1. \end{aligned}$$

Again from Eq. (4.93):

$$\begin{aligned} g_{22} &= \frac{\partial x^1}{\partial \bar{x}^2}\frac{\partial x^1}{\partial \bar{x}^2} + \frac{\partial x^2}{\partial \bar{x}^2}\frac{\partial x^2}{\partial \bar{x}^2} + \frac{\partial x^3}{\partial \bar{x}^2}\frac{\partial x^3}{\partial \bar{x}^2} \\ &= \left(\frac{\partial x}{\partial \theta}\right)^2 + \left(\frac{\partial y}{\partial \theta}\right)^2 + \left(\frac{\partial z}{\partial \theta}\right)^2, \text{ [using (4.97)]} \\ &= (\cos\theta)^2 + (\sin\theta)^2 + (0)^2, \text{ [using (4.96)]} \\ &= r^2. \end{aligned}$$

Again from Eq. (4.93):

$$\begin{aligned} g_{33} &= \frac{\partial x^1}{\partial \bar{x}^3}\frac{\partial x^1}{\partial \bar{x}^3} + \frac{\partial x^2}{\partial \bar{x}^3}\frac{\partial x^2}{\partial \bar{x}^3} + \frac{\partial x^3}{\partial \bar{x}^3}\frac{\partial x^3}{\partial \bar{x}^3} \\ &= \left(\frac{\partial x}{\partial z}\right)^2 + \left(\frac{\partial y}{\partial z}\right)^2 + \left(\frac{\partial z}{\partial z}\right)^2, \text{ [using (4.97)]} \\ &= (0)^2 + (0)^2 + (1)^2, \text{ [using (4.96)]} \\ &= 1. \end{aligned}$$

Again from Eq. (4.93):

$$g_{12} = \frac{\partial x^1}{\partial \bar{x}^1}\frac{\partial x^1}{\partial \bar{x}^2} + \frac{\partial x^2}{\partial \bar{x}^1}\frac{\partial x^2}{\partial \bar{x}^2} + \frac{\partial x^3}{\partial \bar{x}^1}\frac{\partial x^3}{\partial \bar{x}^2}$$
$$= \left(\frac{\partial x}{\partial r}\right)\left(\frac{\partial x}{\partial \theta}\right) + \left(\frac{\partial y}{\partial r}\right)\left(\frac{\partial y}{\partial \theta}\right) + \left(\frac{\partial z}{\partial r}\right)\left(\frac{\partial z}{\partial \theta}\right), \text{ [using (4.95)]}$$
$$= (\sin\theta\cos\phi)(r\cos\theta\cos\phi) + (\sin\theta\sin\phi)(r\cos\theta\sin\phi) + (\cos\theta)(-r\sin\theta)$$
$$= r\sin\theta\cos\theta(\cos^2\phi + \sin^2\phi) - r\sin\theta\cos\theta$$
$$= 0$$

Similarly, $g_{13} = g_{31} = 0$; $g_{23} = g_{32} = 0$. In general $g_{\mu\nu} = 0$, for $\mu \neq \nu$.

Thus

$$ds^2 = g_{\bar{\mu}\bar{\nu}} d\bar{x}^{\mu} d\bar{x}^{\nu}$$
$$= \sum_{\mu=1}^{3} g_{\bar{\mu}\bar{\nu}} d\bar{x}^{\mu} d\bar{x}^{\mu}, \text{ [since, } g_{\mu\nu} = 0, \text{ for } \mu \neq \nu]$$
$$= g_{\bar{1}\bar{1}}(d\bar{x}^1)^2 + g_{\bar{2}\bar{2}}(d\bar{x}^2)^2 + g_{\bar{3}\bar{3}}(d\bar{x}^3)^2$$
$$= (1)(dr)^2 + (r^2)(d\theta)^2 + (1)(dz)^2$$
$$= dr^2 + r^2 d\theta^2 + dz^2$$

4.13 Change of basis: relation between coordinate basis vectors

Change of basis

Any n linearly independent vectors in n-dimensional space (R^n) form a basis in R^n. Let ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$) and ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$) be two sets of bases in 3-dimensional space (R^3). For example let's take

$$\mathbf{e1} = \mathbf{i}_1 + 2\mathbf{i}_2,\ \mathbf{e}_2 = \mathbf{i}_1 + 2\mathbf{i}_2 + \mathbf{i}_3,\ \mathbf{e}_3 = -\mathbf{i}_2 + \mathbf{i}_3 \tag{4.98}$$

and

$$\mathbf{u}_1 = \mathbf{i}_1 + \mathbf{i}_2,\ \mathbf{u}_2 = \mathbf{i}_2 + 2\mathbf{i}_3,\ \mathbf{e}_3 = 2\mathbf{i}_1 + \mathbf{i}_2 - \mathbf{i}_3, \tag{4.99}$$

where ($\mathbf{i}_1$, $\mathbf{i}_2$, $\mathbf{i}_3$) is the standard basis in R^3. These two sets form a basis in $R3$ since the vectors in each set are linearly independent. Now,

$$(\mathbf{e_1} \times \mathbf{e_2})\cdot\mathbf{e_3} = \det\begin{bmatrix} 1 & 2 & 0 \\ 1 & 2 & 1 \\ 0 & -1 & 1 \end{bmatrix} = 1 \neq 0.$$

$$(\mathbf{u_1} \times \mathbf{u_2})\cdot\mathbf{u_3} = \det\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 2 & 1 & -1 \end{bmatrix} = 1 \neq 0.$$

If we want to change from the old basis ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$) to the new basis ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$), we can express each of the new basis vectors to the old ones as follows:

$$(\mathbf{u_i} = \alpha_\mathbf{i}^1\mathbf{e_1} + \alpha_\mathbf{i}^2\mathbf{e_2} + \alpha_\mathbf{i}^3\mathbf{e_3}, i = 1,2,3 \qquad (4.100)$$

where

$$\alpha = \det\begin{bmatrix} \alpha_1^1 & \alpha_1^2 & \alpha_1^3 \\ \alpha_2^1 & \alpha_2^2 & \alpha_2^3 \\ \alpha_3^1 & \alpha_3^2 & \alpha_3^3 \end{bmatrix} \qquad (4.101)$$

is the matrix of the coefficients of direct transformation from the old to new basis. Its i^{th} row is the coordinates of $\mathbf{u}_i$ in the old basis ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$). To find α we have to solve Eq.(4.100) for given ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$) and ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$). For example, for the bases given by Eq.(4.98) & (4.99), relation (4.100) becomes

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \alpha_1^1\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \alpha_1^2\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \alpha_1^3\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \qquad (4.102)$$

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \alpha_1^1\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \alpha_2^2\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \alpha_2^3\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \qquad (4.103)$$

$$\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \alpha_3^1\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \alpha_3^2\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \alpha_3^3\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \qquad (4.104)$$

Rather than solving each of the above system of equations separately, we can combine relations 4.102 to 4.104 to a single system

$$\mathbf{U} = \alpha\mathbf{E} \tag{4.105}$$

where **U** and **E** are matrices whose rows are **ui** and **ei** respectively. In our example we obtain

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 2 & 1 & -1 \end{bmatrix} = \begin{bmatrix} \alpha_1^1 & \alpha_1^2 & \alpha_1^3 \\ \alpha_2^1 & \alpha_2^2 & \alpha_2^3 \\ \alpha_3^1 & \alpha_3^2 & \alpha_3^3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 1 & 2 & 1 \\ 0 & -1 & 1 \end{bmatrix} \tag{4.106}$$

Solving this linear system we find

$$\alpha = \mathbf{UE^{-1}} \tag{4.107}$$

4.14 Change of tensor components under change of coordinate system

Let us consider a tensor A^i_{jk} defined in terms of the coordinates x^i, i = 1,2,3. This tensor can be transformed to $\bar{A}^p_{qr}$ in a new set of coordinates, say $\bar{x}^i, i = 1,2,3$ according to

$$\bar{A}^p_{qr} = \frac{\partial \bar{x}^p}{\partial x^i}\frac{\partial x^j}{\partial \bar{x}^q}\frac{\partial x^k}{\partial \bar{x}^r}A^i_{jk}. \tag{4.108}$$

It is to be remembered that the position of the indices p,q,r on the lefthand side of the transformation are the same as those on the right-hand side. Since p. q, r is associated with □Px coordinates and since i, j, k are associated respectively with p, q, r.

Exercises

1. A quantity$M(p, q, r, s)$ which is a function of coordinates xi transforms to another coordinate system $\bar{x}^i$ according to the rule

$$\bar{M}(i,j,k,l) = \frac{\partial x^p}{\partial \bar{x}^i}\frac{\partial \bar{x}^j}{\partial x^q}\frac{\partial \bar{x}^k}{\partial x^r}\frac{\partial \bar{x}^l}{\partial x^s}M(p,q,r,s) \tag{4.109}$$

(a) Is the given quantity a tensor?

(b) If yes, express thee tensor in suitable notation and

(c) Give the contravariant and covariant order and rank.

2. A covariant tensor has components xy, $2y - z^2$, xz in rectangular coordinates. Find its covariant components in spherical coordinates.

Solutions

1. (a) Yes, the given quantity is a tensor.

(b) M_p^{qrs}.

(c) Contravariant order: 3, covariant order: 1, and rank = 3+1 = 4.

2. Let A_k denotes covariant components in rectangular coordinates

$x_1 = x$, $x_2 = y$, $x_3 = z$. Then

$A_1 = x_1 x_2$

$A_2 = 2x_2 - x_3^2$

$A_3 = x_1 x_3$

Let $\bar{A}_p$ denotes covariant components in spherical coordinates $\bar{x}_1 = r$, $\bar{x}_2 = \theta$, $\bar{x}_3 = \phi$. Then $\bar{A}_p = \frac{\partial x^k}{\partial \bar{x}^p} A_k$. The transformation equations between coordinate systems are

$x_1 = \bar{x}_1 \sin \bar{x}_2 \cos \bar{x}_3$

$x_2 = \bar{x}_1 \sin \bar{x}_2 \sin \bar{x}_3$

$x_3 = \bar{x}_1 \cos \bar{x}_2$

Hence, the covariant components of the given tensor in spherical coordinates are

$$\bar{A}_1 = \frac{\partial x^1}{\partial \bar{x}^1} A_1 + \frac{\partial x^2}{\partial \bar{x}^1} A_2 + \frac{\partial x^3}{\partial \bar{x}^1} A_3$$

$$= (\sin \bar{x}_2 \cos \bar{x}_3)(x_1 x_2) + (\sin \bar{x}_2 \sin \bar{x}_3)(2x_2 - x_3^2) + (\cos \bar{x}_2)(x_1 x_3)$$

$$= (\sin\theta \cos\phi)(r^2 \sin^2\theta \sin\phi \cos\phi) + (\sin\theta \sin\phi)$$

$$\times (2r \sin\theta \sin\phi - r^2 \cos^2\theta) + (\cos\theta)(r^2 \sin\theta \cos\theta \cos\phi)$$

$$\bar{A}_2 = \frac{\partial x^1}{\partial \bar{x}^2} A_1 + \frac{\partial x^2}{\partial \bar{x}^2} A_2 + \frac{\partial x^3}{\partial \bar{x}^2} A_3$$

$$= (r\cos\bar{x}_2 \cos\bar{x}_3)(x_1 x_2) + (r\cos\bar{x}_2 \sin\bar{x}_3)(2x_2 - x_3^2) + (-r\sin\bar{x}_2)(x_1 x_3)$$

$$= (r\cos\theta\cos\phi)(r^2 \sin^2\theta \sin\phi\cos\phi) + (r\cos\theta\sin\phi)$$

$$\times (2r\sin\theta\sin\phi - r^2\cos^2\theta) + (-r\sin\theta)(r^2\sin\theta\cos\theta\cos\phi)$$

$$\bar{A}_3 = \frac{\partial x^1}{\partial \bar{x}^3} A_1 + \frac{\partial x^2}{\partial \bar{x}^3} A_2 + \frac{\partial x^3}{\partial \bar{x}^3} A_3$$

$$= (-r\sin\bar{x}_2 \sin\bar{x}_3)(x_1 x_2) + (r\sin\bar{x}_2 \cos\bar{x}_3)(2x_2 - x_3^2) + (0)(x_1 x_3)$$

$$= (-r\sin\theta\sin\phi)(r^2\sin^2\theta\sin\phi\cos\phi) + (r\sin\theta\cos\phi)$$

$$\times (2r\sin\theta\sin\phi - r^2\cos^2\theta) + (0)(r^2\sin\theta\cos\theta\cos\phi)$$

4.15 Example: Inertial coordinates and bases in Minkowski space

Minkowski space in mathematical physics is the mathematical setting in which Einstein's theory of special relativity is most conveniently formulated. In this setting, the three ordinary dimensions of space are combined with a single dimension of time to form a fourdimensional manifold for representing spacetime. The name Minkowski space (or Minkowski spacetime) is named after the mathematician Hermann Minkowski.

In 1905-06 Henri Poincar′(e) showed that by taking time to be an imaginary fourth spacetime coordinate *ict*, where *c* is the speed of light and $i\left(=\sqrt{1}\right)$ is the imaginary unit, a Lorentz transformation can formally be regarded as a rotation of coordinates in a fourdimensional space with three real coordinates representing space, and one imaginary coordinate representing time, as the fourth dimension. In physical spacetime special relativity stipulates that the quantity: $-t^2 + x^2 + y^2 + z^2$ is invariant under coordinate changes from one in ertial frame to another, i. e. under Lorentz transformations. Here the speed of light *c* is, following Poincaré, set to unity. In the space suggested by Poincaré, in which spacetime is represented by coordinates $(t, x, y, z) \leftrightarrow (x, y, z, it)$, known as coordinate space, Lorentz transformations appear as ordinary rotations preserving the quadratic form $x^2 + y^2 + z^2 + t^2$ on coordinate space. The naming and ordering of coordinates, with the same labels for space coordinates, but with the imaginary time coordinate as the fourth coordinate, is conventional. The above expression, while making the former expression more familiar, may potentially be confusing because it is not the same t that appears in the latter (time coordinate) as in the former (time itself in some inertial system as measured by clocks stationary in that system).

Standard basis in Minkowski space

In his 1908 "Space and Time" lecture, Minkowski gave an idea of using a real time coordinate instead of an imaginary one, representing the four variables (x, y, z, t) of space and time in coordinate form in a four dimensional real vector space. Points in this space correspond to events in spacetime. In this space, there is a defined light-cone associated with each point, and events not on the light-cone are classified by their relation to the apex as spacelike or timelike. It is principally this view of spacetime that is current nowadays, although the older view involving imaginary time has also influenced special relativity. The vector space nature of Minkowski space allows for the canonical identification of vectors in tangent spaces at points (events) with vectors (points, events) in Minkowski space itself. They can be expressed formally in Cartesian coordinates as $(x^0, x^1, x^2, x^3) \leftrightarrow x^0 e_0|_p + x^1 e_1|_p + x^2 e_2|_p + x^2 e_2|_p + x^3 e_3|_p \leftrightarrow x^0 e_0|_q + x^1 e_1|_q + x^2 e_2|_q + x^3 e_3|_q$, with basis vectors in the tangent spaces defined by

$$e_\mu|_p = \frac{\partial}{\partial x^\mu}|_q, \; or, \; e_0|_p = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, etc \tag{4.110}$$

where p and q refer to any two events. The first identification is the canonical identification of vectors in the tangent space at any point with vectors in the space itself. The appearance of basis vectors in tangent spaces as first order differential operators is due to this identification.

A standard basis for Minkowski space is a set of four mutually orthogonal vectors $\mathbf{e}_0$, $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ such that $-(e_0)^2 = (e_1)^2 = (e_2)^2 = (e_3)^2 = 1$. These conditions can be written compactly in the following form:

$$< e_\mu, e_\nu > = \eta_{\mu\nu}, \; [\mu, \nu = 0,1,2,3] \tag{4.111}$$

where η is given by

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (4.112)$$

The above tensor (η) is frequently called the "Minkowski tensor". Relative to a standard basis, the components of a vector v are written (v^0, v^1, v^2, v^3), which in Einstein notation is written as $v = v^\mu e_\mu$. The component v^0 is called the timelike component of v while the other three components are called the spatial components.

4.16 Lorentz transformations as coordinate transformations

In physics, the Lorentz transformations are a one-parameter family of linear transformations from a coordinate frame in spacetime to another frame that moves at a constant velocity (the parameter) relative to the former. The transformations are named after the Dutch physicist Hendrik Lorentz. The respective inverse transformation is then parameterized by the negative of this velocity.

The most common form of the transformation, parametrized by the real constant v, representing a velocity confined to the x-direction, is expressed as

$$\left.\begin{aligned} t' &= \gamma(t - \frac{v_x}{c^2} \\ x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \end{aligned}\right\} \qquad (4.113)$$

where (t, x, y, z) and (t', x', y', z') are the coordinates of an event in two frames, where the primed frame is seen from the unprimed frame as moving with speed v along the x-axis, c is the speed of light, and $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ is the Lorentz factor. When speed v is significantly lower than c, the factor is negligible, but as v approaches c, there is a significant effect. The value of v cannot exceed c, in the present understanding.

4.17 Elelctro-magnetic tensor and change in its components under Lorentz transformations

Introduction

In electromagnetic theory, the electromagnetic tensor or electromagnetic field tensor (sometimes called the field strength tensor, Faraday tensor, or Maxwell bivector) is a mathematical object that describes the electromagnetic field in spacetime. The field tensor was first used after the four-dimensional tensor formulation of special relativity was introduced by Hermann Minkowski. The tensor allows related physical laws to be written very concisely.

Derivation of EM filed tensor using Lorentz gauge

It is possible to combine three-components of each of electric field $\vec{E}$ and magnetic field $\vec{B}$ to construct a second rank tensor $F^{\mu\nu}$ in the fourdimensional space-time. In the Lorentz gauge, $\vec{E}$ and $\vec{B}$ are given in terms of the scalar and vector potentials as

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} - \nabla\phi \tag{4.114}$$

$$\vec{B} = \nabla \times \vec{A} \tag{4.115}$$

The x-components of the above equations are

$$E_x = -\frac{1}{c}\frac{\partial A_x}{\partial t} - \frac{\partial \phi}{\partial x} = -\partial^0 A^1 - \partial^1 A^0 \tag{4.116}$$

$$B_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = \partial^2 A^3 - \partial^3 A^2 \tag{4.117}$$

The above forms of Ex and Bx prompt that one may define a second rank tensor

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu = \frac{\partial A^\nu}{\partial x_\mu} - \frac{\partial A^\mu}{\partial x_\nu} \tag{4.118}$$

Now,

$$F^{\nu\mu} = \partial^\nu A^\mu - \partial^\mu A^\nu = -(\partial^\mu A^\nu - \partial^\nu A^\mu) = -F^{\mu\nu}, \tag{4.119}$$

so $F^{\mu\nu}$ is an antisymmetric tensor. i.e.,

$$F^{\mu\nu} = -F^{\nu\mu}, \tag{4.120}$$

as a matter of fact diagonal elements of $F^{\mu\nu}$ vanish. If in Eq(4.72) on choses ν = 0, μ = 1, 2, 3, he may check from Eq. (4.70) that

$$F^{10} = E_x,\ F^{20} = E_y,\ F^{30} = E_z \tag{4.121}$$

Further if one chooses the values of the index pair ($\mu\nu$) to be (1,2), (2,3), (3, 1), he may see from Eq. (4.71) that

$$F^{12} = -B_z, F^{23} = -B_x, F^{31} = -B_y \tag{4.122}$$

The tensor $F^{\mu\nu}$ then becomes

$$F^{\mu\nu} = \begin{Bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{Bmatrix} \tag{4.123}$$

The tensor represented by Eq.(4.77) is called electromagnetic field tensor. Eq.(4.77) gives $F^{\mu\nu}$ explicitly in terms of $\vec{E}$ and $\vec{B}$ and Eq.(4.72) represented $F^{\mu\nu}$ in terms of the scalar and vector potentials ϕ and $\vec{A}$ respectively.

Lorentz transformation of EM Field tensor

We can use the usual tensor transformation rules to see how the electric and magnetic fields transform under a Lorentz transformation. We get

$$F^{\mu\nu} = \frac{\partial x'^{\mu}}{\partial x^i}\frac{\partial x'^{\nu}}{\partial x^j}F^{ij} \tag{4.124}$$

$$= \Lambda_i^{\mu}\Lambda_j^{\nu}F^{ij} \tag{4.125}$$

where the Lorentz transformation matrix is

$$\Lambda_i^{\mu} = \begin{Bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Bmatrix} \tag{4.126}$$

The above transformation (4.75) can be written as a matrix equation of the form

$$F' = \Lambda F \Lambda^T \tag{4.127}$$

The first product

$$\Lambda F = \begin{Bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Bmatrix} \begin{Bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{Bmatrix} \tag{4.128}$$

$$= \begin{Bmatrix} -\gamma\beta E_x & -\gamma E_x & -\gamma Ey + \gamma\beta B_z & -\gamma E_z - \gamma\beta B_y \\ \gamma E_x & \gamma\beta E_x & \gamma\beta E_y - \gamma B_z & \gamma\beta E_z + \gamma B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{Bmatrix} \tag{4.129}$$

The final product

$$F' = \Lambda F \Lambda^T$$

$$= \begin{Bmatrix} -\gamma\beta E_x & -\gamma E_x & -\gamma Ey + \gamma\beta B_z & -\gamma E_z - \gamma\beta B_y \\ \gamma E_x & \gamma\beta E_x & \gamma\beta E_y - \gamma B_z & \gamma\beta E_z + \gamma B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{Bmatrix}$$

$$\times \begin{Bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Bmatrix}$$

$$= \begin{Bmatrix} 0 & -E_x & -\gamma Ey + \gamma\beta B_z & -\gamma E_z - \gamma\beta B_y \\ E_x & 0 & \gamma\beta E_y - \gamma B_z & \gamma\beta E_z + \gamma B_y \\ \gamma E_y - \gamma\beta B_z & -\gamma\beta E_y + \gamma B_z & 0 & -B_x \\ \gamma E_z + \gamma\beta B_y & -\gamma\beta E_z - \gamma B_y & B_x & 0 \end{Bmatrix} \tag{4.130}$$

Again in the rotated frame

$$F' = \begin{Bmatrix} 0 & -E'_x & -E'_y & -E'_z \\ E'_x & 0 & -B'_z & B'_y \\ E'_y & B'_z & 0 & -B'_x \\ E'_z & -B'_y & B'_x & 0 \end{Bmatrix} \tag{4.131}$$

Comparing (4.120) and (4.121) element by element we have the components of EM field tensor under Lorentz transfomation

$$E'_x = E_x \tag{4.132}$$

$$E'_y = \gamma E_y - \gamma\beta B_z \tag{4.133}$$

$$E'_z = \gamma E_z + \gamma\beta B_y \tag{4.134}$$

$$B'_x = B_x \tag{4.135}$$

$$B'_y = \gamma\beta E_z + \gamma B_y \tag{4.136}$$

$$B'_z = -\gamma\beta E_y + \gamma B_z \tag{4.137}$$

It is to be noted that, unlike lengths, the components of E and B in the direction of motion are unchanged, while those perpendicular to the motion are altered.

Chapter-end exercise

1. What is the Levi-Civita symbol? What is contraction applied to tensors?
2. What do you mean by contravariant, co-variant, and mixed tensors? Prove that velocity and acceleration are contravariant and the gradient of a field is a covariant tensor.
3. Show that any tensor of rank 2 can be expressed as the sum of symmetric and antisymmetric tensors of rank 2.
4. Construct a scalar from the tensor V^{ij}_{kl}.

References

- Complex Variables and Applications, *J. W. Brown & R. V. Churchill*, 9th Edition, Mc Graw Hill, 2014
- Mathematical Methods for Physicists, *Arfken & Webber*, 6th Edition, Elsevier, 2005
- Schaum's Outlines: Complex Variables, *M. R. Spiegel, S. Lipschutz, J. J. Shiller, D. Spellman*, 2nd Edition, Mc Graw Hill, 2009

NOTE