First Semester B.Sc Physics

V.N.Purushothaman

23/11/2012

Contents

Syllabus

Core Course I
PH1 B01: METHODOLOGY OF SCIENCE AND PHYSICS- 36 hours (Credit - 2)
Part A: Methodology And Perspectives Of Sciences (10Hours)

Unit I - Science and Science Studies: Types of knowledge: Practical, Theoretical, and Scientific knowledge, Information. What is Science; what is not science; laws of science. Basis for scientific laws and factual truths. Science as a human activity, scientific temper, empiricism, vocabulary of science, science disciplines. Revolution in science and Technology. *Unit II - Methods and tools of science:* Hypothesis: Theories and laws in science. Observations, Evidences and proofs. Posing a question; Formulation of hypothesis; Hypothetico-deductive model, Inductive model. Significance of verification (Proving), Corroboration and falsification (disproving), Auxiliary hypothesis, Ad-hoc hypothesis. Revision of scientific theories and laws, Importance of models, Simulations and virtual testing, Mathematical methods vs. scientific methods. Significance of Peer Review.
Reference Books:

1. Gieryn, T F. Cultural Boundaries of Science., Univ. of Chicago Press, 1999
2. Collins H. and T Pinch., The Golem: What Everyone Should Know About Science., Cambridge Uni. Press, 1993
3. Hewitt, Paul G, Suzanne Lyons, John A. Suchocki & Jennifer Yeh, Conceptual Integrated Science. Addison-Wesley, 2007
4. Newton R G. The Truth of Science: New Delhi, 2nd edition
5. Bass, Joel E and et. al. Methods for Teaching Science as Inquiry, Allyn & Bacon, 2009

Part B: Methodology and Perspectives of Physics (12Hours)

(All topics in this part require qualitative study only, derivations are not required) What does physics deal with? - brief history of physics during the last century-the inconsistency between experiments and theories- Birth of new science concepts -Quantum concepts-Black body radiation, Photoelectric effect, X-rays, Compton effect, De Broglie waves, Sections 2.2, 2.3, 2.5, 2.7, 3.1, of Arthur Beisser)
Relativity-Special relativity, Time dilation, Length contraction, Twin paradox *(Sections 1.1, 1.2, 1.4, 1.5 of Arthur Beisser)*
Laser- Concepts of ordinary and monochromatic light, Coherent and incoherent light, Spontaneous and stimulated emission, Metastable state, pumping and population inversion. (Basic ideas only Section 4.9 of Arthur Beisser)
Design of an experiment , experimentation , Observation, data collection: Interaction between physics and technology.
References:
1. Concepts of Modern physics- Arthur Beisser
2. A brief history and philosophy of Physics - Alan J. Slavin- http:/ www.trentu. Ca/ academic / history- 895 .html
3.The inspiring History of Physics in the Last One Hundred Years : Retrospect and prospect Prof. Dr-lng.Lu Yongxiang http :// www.twas .org.cn/twas/proLu.asp

Part C - Mathematical Methods in Physics (14 Hours)

Vector Analysis: - Vector Operations - Vector Algebra - Component form - How vectors transform, Applications of vectors in Physics. Differential Calculus: - The operator ∇ - Gradient, Divergence, Curl - Physical interpretation - Product rules of ∇ - Second derivatives.
Integral Calculus: - Line integral, surface integral and volume integral - Fundamental theorem of Gradients - Gauss's Divergence Theorem (Statement only)- The fundamental theorem of curl - Stoke's theorem(Statement only). Divergence less and curlless fields.
Curvilinear co-ordinates: - Spherical polar coordinates - cylindrical coordinates(Basic ideas).

Matrices: - Basic ideas of matrices - addition, subtraction, scalar multiplication, Transpose of a matrix, conjugate of a matrix, diagonal matrix - Representation of vectors as column matrix - Determinants - Cramer's rule - Eigen Values and Eigen Vectors - Hermitian Matrix,Unitary Matrix.

References:

1. Introduction to electrodynamics - David J . Griffiths, Prentice Hall India Pvt. Ltd., Chapter - 1 2. Mathematical Physics - Satya Prakash, Sultan Chand & Sons, New Delhi

axiom
empirical
empirical

1 Methodology And Perspectives Of Sciences

Science and Science Studies

Types of knowledge:

Knowledge may be defined as belief which is in agreement with the facts. For example, when a man is hungry, he believes that taking food will reduce his hunger. This is an agreed fact. His belief is thus knowledge. Depending on the source, knowledge is classified into three types.

Theoretical knowledge: It is knowledge acquired by rational thinking or intellectual insight. eg. Einstein's assertion that velocity of light must be a constant always and everywhere is an intelligent insight. Higher mathematics is pure theoretical knowledge.

Practical knowledge: It is knowledge acquired through experience. eg. Swimming and cycling can be learned only through experience.

Scientific knowledge: It is knowledge accumulated by systematic study and organized by general principles. eg. The observation that heat by itself will never flow from cold body to hot body is later stated as second law of thermodynamics.

Information: It is data that has been verified to be accurate, which increases understanding and decreases uncertainty. The report of sighting of a new star in the horizon is an information as it may explain the unexpected behaviour of nearby stars. The data $log_{10}2 = 0.301029996$ is an information because it establishes the relation between 2 and 0.301029996. Technically information is a sequence of symbols that can be interpreted as a message

Definition of Science: Science is a precise, systematic and objective method of studying nature. It starts with the axiom that there is some form of order in nature. To discover this order, the method involves collection and interpretation of data and finally generalizing into physical principles. Data collection is possible only if direct or indirect observations could be made. For precision of data, these observations must be expressible in terms of natural numbers. Such observations are called measurements. Any branch of knowledge that relies solely on measurements is called empirical. Science is thus empirical.

Laws of science: Laws of science are statements expressing regularities of nature as precisely as possible. If these regularities are observed at all places and all times without exception, it is called a universal law. *(eg.1)*It is impossible to attain absolute zero of temperature by a finite cycle of cooling operations.*(eg.2)* The path of a body falling freely is a geodesic.

If only a certain fraction of cases show regularity, such laws are called statistical laws. *(eg.1)* If a cube is tossed a large number of times N, approximately $N/6$ times it comes down with same face. *(eg.2)* Approximately half the children born each year are girls.

Basis for scientific laws and factual truths: Facts are particular events. For example, the detection of a magnetic field around a current carrying conductor at some instant is a fact. It is a single event and is a reality. It cannot be logically disputed or rejected. Facts are not discovered or created, but simply acknowledged.

A truth is something that must be discovered, or created. If somebody has strong reason to state that there is no upper limit to particle velocity, it is a truth to him . However, if another person has enough reasons to feel the converse, then that is also truth to that person. Truths are subjective while facts are objective. In science facts are often used to substantiate truths. Similarly truths are used to understand certain facts better. But it is wrong to assert a fact as a truth, or *vice versa*. A truth created out of a set of facts is a factual truth. Factual truths are often expressed as scientific laws. *eg.* Newton's law of gravitation is a consequence of the facts that planets circle stars, free-fall is always downward etc.

Science as a human activity: Of all animals only human beings are capable of thinking. This ability helps him to discover patterns and structures in what he observe. These discoveries lead to the creation of laws. The set of natural laws constitute science. Hence science is exclusively a human activity.

Scientific temper: It refers to an attitude which involves the application of logic and the avoidance of preconceived notions. Discussion, argument and analysis are the important features of scientific temper. It admits every point of view from everywhere at all times.

Empiricism: The term empiricism in ancient Greek means 'experience'. It refers to the idea that all concepts originate and justified in experience. It is applicable to things that can be experienced. Put forward by Hempel and developed by David Hume, empiricism argues that as it is impossible to experience causal relations, causality is pure imagination. But it is a fact that causality is indispensable to understand the world around us. Hence explanations based on causality is more acceptable even today.

vocabulary of science: It is a collection of words or phrases or symbols with a well-defined meaning and used in science. The vocabulary is used to express all concepts that are relevant to the understanding of science. A few of them are given below.

1. *Matter* is anything that has mass and take up space
2. *Atom* is defined as a small particle that makes up most types of matter.
3. *Conservation* means no change with time.
4. *Electrons* invisible , negative charged particle usually surround the nucleus of an atom.
5. *Nucleus*is positively charged, central part of an atom.
6. *Protons* are positively charged, particle located in the nucleus of an atom.
7. *Neutrons* are uncharged particles located in the nucleus of an atom.
8. *Element* is matter made of only one kind of atom.
9. *Isotope* are atoms of the same element that have different numbers of neutrons.
10. *Atomic number* tells you the number of protons in nucleus of each atom of
11. element.
12. *Mass number* is the number of protons plus the number of neutrons it contains.
13. *Atomic mass* is weighted average mass of isotopes of an element.
14. *Metal* generally have a shiny or metallic lustre and one good conductors of heat and electricity.
15. *Nonmetal* are elements that are usually dull in appearance.
16. *Metalloid* matter with a fixed composition whose identity can be changed by chemical processes
17. *Compound* is a substance whose smallest unit is made up of an atoms of more than element bounded together.
18. *Mixture* is a combination of compounds and elements that has not formed a new substance and whose proportions can be change without changing the mixture identity.
19. *Solid* is matter with a definite shape and volume.
20. *Liquid* is matter that has a definite volume but no shape.
21. *Gas* is matter that does not have a definite shape or volume.
22. *Viscosity* results from the coherent force between the particles of a fluid.

23. *Surface tension* causes the liquid to act as if a thin film were stretched across its surface.

24. *Thermal energy* is an extensive property, depends on the number of particles in a substance as well as the amount of energy each particle has.

25. *Temperature* an intensive property of substance.

26. *Heat* is disordered energy.

27. *Melting* The process of a solid becoming a liquid.

28. *Vaporization* The process of a liquid becoming a gas.

29. *Condensation*The process of a gas becoming a liquid.

30. *Pressure* is equal to force exerted per unit area.

31. *Buoyant* force is equal to the weight of fluid displaced.

Science disciplines: It is any particular branch of scientific knowledge. *eg.* physical science, biological science *etc* There is a basis for the development of different disciplines.

Science is a method to study nature. Nature consists of living and non-living objects. The study of living things gave rise to biological sciences while that of non-living led to physical sciences. Biological science then branched out to immobile (botany) and self-mobile(zoology)living systems. Then another branch called genetics is developed to study life as a physical phenomenon. There are a large number of such branches of biology.

In physical science, the study of fundamental nature of objects and interactions between them paved the way for Physics. The properties, structures and the permanent changes that occur to them developed as chemistry. In a similar way sub-disciplines like particle physics, nuclear physics, molecular dynamics, quantum mechanics, statistical physics, inorganic chemistry, organic chemistry, physical chemistry *etc* were also developed. The development of new disciplines is a continuous process as our understanding of nature increases continuously. It is also found that certain aspects of different disciplines merge to give rise to interdisciplinary subjects. Environmental science, meteorology, atmospheric science computer science *etc* are examples.

Revolution in Science : The scientific revolution refers to a paradigm shift with the introduction of new ideas in natural sciences. For example, with the publication of *'On the Revolutions of the Heavenly Spheres'* by Nicolaus Copernicus and in the 16^{th} century, the idea that planets revolve with a star as centre and not *vice versa* is established. Similarly *'On the Fabric of the Human body'* by Andreas Vesalius firmly established the flow of blood in the human body. Started in Europe, it continued upto the 19^{th} century. Thermodynamics, electricity, magnetism, mechanics *etc* were developed by Joule, Faraday, D'Alembert, Kepler, Rene Descartes, Isaac Newton *etc*. Similar developments occurred in other branches of science like medicine, human anatomy, organic chemistry *etc*.

The next important changes in the concept of nature occurred in the beginning of 20^{th} century with the advent of quantum theory by Max Planck and theory of relativity by Albert Einstein. The quantization of gravity may be the next revolutionary change in our perception of the microscopic and macroscopic nature.

Revolution in Technology: A scientific revolution is usually followed by a revolution in Technology. With the development of thermodynamics and the idea of a heat engine by Sadi Carnot, James Watt developed the Steam engine. This engine is the key in automation, transportation, and running machines in industry. With the discovery of petroleum, internal combustion engines were developed to run light vehicles and industrial machines. Industrial revolution was a consequence of these technologies.

In the 20^{th} century, the application of quantum mechanics to crystalline solids led to the development of semiconductor electronics. The microprocessors and computer hardware are made of electronic components burned into semiconductor chips. The use of these devices are so widespread that one cannot mention even a single sphere of life electronics is not used. A large number of new drugs are the application of genetic engineering. The green revolution which substantially increased food production is a consequence of genetic manipulation.

2 Unit II - Methods and tools of science

Methods of scientific enquiry

There are four methods usually employed. They are inductive, deductive, constructive and hypothetico-deductive method.

Inductive method: In this method, facts are collected from observations. They are then analysed to find out possible generalizations. It is thus a method of generating theory from facts. Two ideas are implicit here. (1)uniformity of nature and (2)causality. Uniformity of nature means that if an event occurs here under a certain condition, it will recur at all places and all times if the same condition persists. Causality means every event has a non-empty set of causes. Nothing happens by chance. In non-relativistic physics, mass is a conserved quantity as its creation or destruction is never observed. This is an example of inductive method.

Deductive Method: If conclusions about a particular case are reached by mere logical reasoning from some general principle, it is deductive method. The conclusion that planets must be confined to fixed orbits due to gravitation is an example of deductive method. Usually it is more efficient than inductive method.

empirical
hypothe
hypothesi
hypothesi
hypothe
hypothesi

Constructive Method: If new principles are derived from existing principles without empirical verification, it is called constructive method. It is similar to deductive method and is used in mathematics. For example most of the theorems of plane geometry are derived from Euclid's axioms.

hypothetico-deductive method: In this method, a minimum number of principles are initially assumed and a few conclusions are derived. These conclusions are then compared with observations. If they do not match, the initial principles are modified. If no match is obtained, it is usually rejected. *eg.* Kinetic theory of gases assumes intermolecular force and molecular volume as negligible. But at high pressures and low temperatures, they become significant. Perfect gas equation is to be replaced with Van der Wall's equation.

hypothesis

It is a set of propositions used as an explanation for the occurrence of a specified set of phenomena. It postulates a possible relationship between two or more phenomena or variables.If it is employed as a temporary concept to start an investigation, it is called a working hypothesis. If it is accepted as the most probable idea that explains the established facts, it becomes a basic principle. The hypothesis must satisfy the following conditions.

- It should simplify the problem.
- It should give an acceptable explanation of the phenomenon.
- It must be verifiable.
- It must be formulated in simple, understandable terms.

Formulation of hypothesis:

A hypothesis is formulated either to give an explanation of some new phenomenon or to solve some problem. The formulation usually involves the following steps.

- Identifying the independent and dependent variables to be studied.
- Specifying the nature of the relationship that exists between these variables.
- Simple and parsimonious.

hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
axiom
hypothesis
empirical
hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
hypothesis
empirical
hypothesis

- If more than one hypothesis is needed, It is better to have several simple ones than one complicated hypothesis.
- It must be independent of any specific measurement methods.
- It must be falsifiable and testable.

Auxiliary hypothesis: During testing of a hypothesis, it is often necessary to introduce a new hypothesis to support the implications of results of testing. This new hypothesis is called an auxiliary hypothesis. *eg.* When dispersion was discovered, appearance of nearly straight rainbow was expected. When rainbow is found to be semicircular, spherical nature of horizon is taken as an auxiliary hypothesis.

Ad-hoc hypothesis: It is an auxiliary hypothesis used to defend a main hypothesis against a negative result. *eg*To make Maxwell's equation galilean invariant, length contraction along direction of motion was assumed.

Observations, Evidences and proofs: An observation is a datum collected using our senses (direct observation) or using instruments (indirect observation). If the object observed is nature and observation is as objective as possible, it is scientific observation. eg. Observation of an accelerating vehicle(direct), the appearance of bacteria in a tissue in its microscopic image(indirect).

Scientific evidence is anything that proves or disproves the truth of an axiom or hypothesis . Evidence is collected by empirical means or properly collected data in accordance with scientific validity.

A proof is a convincing evidence beyond any doubt everywhere and always. A proved hypothesis is a theorem. This is true in mathematics. Since scientific hypotheses are falsifiable, scientific proof is a myth.

Significance of Verification, Corroboration and Falsification: Verification of a hypothesis refers to testing the truth of it. If observed facts agree with the predictions of a hypothesis, that hypothesis is assumed true. A hypothesis that unlike charges attract can be verified by placing unlike charges near to each other. It is direct verification. But the idea that gases are made up of molecules cannot be verified directly. It is then deduced that molecules hitting the container should cause pressure. Pressure is then observed. This is indirect verification.

Corroboration refers to affirmation of a theory with the help of new evidence from different sources. The wave nature of light proposed by Huygenes to explain its spreading is corroborated with the discovery of diffraction and polarization.

A hypothesis is falsifiable if it is testable by empirical means. It means that if a hypothesis is false, then observation will show that is is false at some point. It conforms to the standards of scientific method.

hypothesi
hypothesi
hypothesi

Revision of scientific theories and laws: Scientific theories are never proved but confirmed. If a new experimental evidence is in conflict with current theory, the theory is modified accordingly. If after a few such efforts the new evidence cannot be explained, then the theory is rejected altogether. *eg.* The stability of nucleus against the repulsive proton-proton force led to the hypothesis of an attractive square-well nuclear force. The stability of spin-1 deuteron forced us to modify nuclear force by adding a spin dependence.

Importance of models: A model in science is a representation of an object or event showing its essential characteristics. They are developed either to formulate a hypothesis or when it is known that a hypothesis has only limited validity. The scientific method often involves procedures for the construction and verification of models. For example, we do not know which are the complete set physical variables on which our climate depends. So we choose atmospheric pressure, temperature, humidity, wind speed etc and establish relations between them using principles of fluid dynamics. Using these relations, we compute the climate of past which is known to us. If the model agrees with past observations, it is used to predict future climate.

An example where models are used when has limited validity is nuclear models. To explain binding energy and nuclear fission, liquid drop model is introduced. When magic numbers were discovered (N,Z=2,8,20,28,50,82,126), a shell model for the nucleus is assumed. With the discovery of vibration and rotation spectrum, a collective nuclear model was put forward. All these are models of the same nucleus. Each model is having only limited validity.

Modelling is imperative in the following situations

- The size is too small (subatomic) or too large (cosmic).
- Too far away, too back in history or too remote in future
- Too complicated to study in real world situations, like the working of our brain.

Simulations and virtual testing: Simulation refers to the representation of the structure and behaviour of one system through the use of another system, usually a computer program designed for the purpose. It is a sort of dynamic modelling. The behaviour of a nucleus in which a projected proton gets embedded can be simulated if we know the approximate form of nuclear potential.

Virtual testing is defined as developing and debugging a test for the simulation models of a system. It is a method for testing without a physical device. The ECG is an example of virtual test. Here the signals received from are heart are used to simulate its behaviour using a computer.

Mathematical methods vs. scientific methods: Mathematics provides correct proofs of true theorems. There are standards of correctness and clear grounds for the empirical

idea of truth. The only requirement for a mathematical theorem to be true is that its proof must be logically consistent with the axioms with which the proof began. It need not conform to the structure of reality.

Scientific methods are fundamentally empirical. Whatever result you get, must be consistent with reality. Logical consistency is often compromised in favour of reality. For example, imagine some physical quantity $p \to 0$ and another quantity $q = 1/p$ must be finite. What physicists do is to replace p with $p + i\eta$ where $i = \sqrt{-1}$ and η a very small quantity. When they complete all their calculations, then the limit $\eta \to 0$ gives them a result without divergent terms.

Significance of Peer Review: Peer review is the evaluation of ones work by other people in the same field. It helps to remove irrelevant data, wrong conclusions and false claims. Its aim is to enhance the quality of the work in that field.

3 Part B: Methodology and Perspectives of Physics

Black body radiation:

The term black-body refers to an object or system which absorbs all radiation incident upon it irrespective of the type of incident radiation. It is a *perfect absorber*. For thermal equilibrium, it should also be a *perfect emitter*. The radiation energy emitted by a black-body is found to be different at different frequencies. Wein empirically found a formula to describe the power E_λ radiated by a black-body at a temperature T k as

$$E_\lambda d\lambda = \frac{A\exp\left(-B/kT\right)}{\lambda^5}d\lambda$$

where A and B are adjustable parameters for curve fitting and $K = 1.381\times10^{-23}Jk^{-1}mole^{-1}$. From this he found that if λ_m represents wavelength having maximum energy at temperature T, then

$$\lambda_m T = constant$$

which is called Wein's displacement law. It is widely used to determine temperatures of distant stars from their spectrum. Later Raleigh showed that E_λ is finite even if $T \to \infty$ which is not acceptable. It is also found that the law does not hold for long wavelengths.

Raleigh and Jeans then derived another energy distribution function using principle of equipartition of energy. Equipartition of energy means energy associated with each degree of freedom of a particle is $kT/2$. Their function is

$$E_\lambda d\lambda = \frac{8\pi kT}{\lambda^4}d\lambda$$

This formula is found to hold only for long wavelengths. As $\lambda \to 0$, $E_\lambda \to \infty$ which is clearly unphysical. It is called *Ultraviolet catastrophe*

Max Planck realized that the problem with the theory of radiation is the implicit assumption that radiation is a continuous flow of energy. He assumed that electromagnetic energy is exchanged between systems as discrete 'packets' which he called quanta. The energy of each quantum is given by $E = h\nu$ where ν is frequency and $h = 6.616 \times 10^{-34}Js$ is called Planck's constant. The resulting energy distribution is given by

$$E_\lambda d\lambda = \frac{8\pi hc\, d\lambda}{\lambda^5}\frac{1}{\exp\left(h\nu/kT\right)-1}$$

For small λ, ν is large as $\nu\lambda = c$, a constant.

$$\frac{1}{\exp\left(h\nu/kT\right)-1} \approx \frac{1}{\exp\left(h\nu/kT\right)} = \exp\left(-h\nu/kT\right)$$

$$E_\lambda d\lambda = \frac{8\pi hc}{\lambda^5}\exp\left(-h\nu/kT\right)d\lambda$$

empirical
black-bod
Wein's displace law
Raleigh
Jeans
Equiparti energy
Ultraviole catastro
Max Plan
quanta

Photoelectric emission
photon
x-ray
Compton effect
Compton wavelength

This is identical to Wein's energy distribution with $A = 8\pi hc$ and $B = h\nu$.
For long wavelengths or small frequencies

$$\frac{1}{\exp{(h\nu/kT)} - 1} \approx \frac{1}{(h\nu/kT)} = \frac{1}{(hc/kT\lambda)} = \frac{kT\lambda}{hc}$$

$$E_\lambda d\lambda = \frac{8\pi hc}{\lambda^5}\frac{kT\lambda}{hc}d\lambda = \frac{8\pi kT}{\lambda^4}d\lambda$$

which is Rayleigh-Jeans law.

Experimental verification of quantum theory

Photoelectric effect, Compton effect and Raman effect provide direct verification of the quantum concept of radiation.

Photoelectric effect: It is observed that electrons are emitted from alkaline metals when they are exposed to electromagnetic waves above a certain frequency. The number of electrons emitted increases with incident intensity. No emission will occur below minimum frequency even if the intensity large. Since electrons need a minimum energy E to get out of the metal, the minimum frequency must correspond to that energy. Intensity cannot affect chance for emission means energy is not absorbed continuously. Photoelectric emission occurs when a quantum of light called photon of frequency ν of required energy $E = h\nu$ is absorbed by an electron of the metal. As number of such photons increases, number of electrons emitted also increases. This explains dependence of number of electrons emitted on light intensity.

Compton effect: When an x-ray photon with energy $h\nu$ collides with a stationary electron, part of the energy and momentum is transferred to the electron. This is known as the Compton effect. Energy and momentum are conserved in this collision. After the collision the photon has energy $h\nu'$ and the electron has acquired a kinetic energy K. Special relativity gives initial photon momentum $|\vec{p}| = h\nu/c = h/\lambda$ and final photon momentum $|\vec{p'}| = h\lambda'/c$. Let the electron momentum be $\vec{q}$. Conservation of energy and momentum gives

$$h\nu = h\nu' + K,\ \vec{p} = \vec{p'} + \vec{q}$$

The change in wavelength due to Compton scattering is given by

$$\lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta)$$

where m_e is electron mass and θ the angle of deviation of the photon. $h/m_e c$ is called the Compton wavelength of the electron. This is a convincing evidence that light is made up of photons having linear momentum.

Raman effect: The change in wavelength of light scattered while passing through a transparent medium is called Raman effect. If scattering increases wavelength, they are called Stokes spectral lines. If scattered wavelength is smaller, they are called anti-stokes lines. The set of new wavelengths forms Raman spectrum which is characteristic of the scattering medium.

De Broglie waves or Matter waves: It is a hypothetical wave associated with the motion of a microscopic particle that explains their diffraction by crystals, small holes and sharp edges. It is postulated by Louis Broglie. The wavelength associated with matter waves is given by $\lambda = h/p$ where p is momentum of particle and h is Planck's constant.

X-rays: Rontgen found that when fast electrons are suddenly stopped, highly penetrating radiations are produced. They are called x-rays. They are widely used in medical diagnosis and treatment of cancers.

Lasers

Ordinary and monochromatic light: Ordinary light consists of all wavelengths of the visible region. eg. Sunlight. Monochromatic light consists of only a single wavelength. There is no perfect monochromatic light from any natural or man-made source. Using proper wavelength filters, light that is approximately monochromatic can be produced.

Coherent and incoherent light: Two beams of light are coherent if the phase difference between them is constant. It is possible only if their wavelengths are equal and the beams originate from the same source at the same time.

LASER: It stands for **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. Electronic transitions between discrete energy levels in atoms emit or absorb visible light. Electrons in some of the excited levels will stay there longer. Such levels are called *metastable states*. Some of the electrons in metastable state return to ground state by emitting a photon. It is called *spontaneous emission*. Rest of the electrons de-excite under the influence of a photon of same frequency. It is called *stimulated emission*. If the number of atoms in the metastable state is large, it is called *population inversion*.To achieve significant population inversion, absorbing medium should be given energy in some form. The process of energising a laser medium to get population inversion is called *optical pumping*. The lasing medium is usually placed between a fully silvered and a partially silvered surfaces called an *optical cavity*. After achieving population inversion, a spontaneously emitted photon induces excited atoms to radiate simultaneously so that it is *coherent*. These photons oscillate in the medium between

Raman ef
Stokes sp lines
anti-stoke
Louis Bro
Planck's c
Rontgen
Monochro light
coherent
LASER
metastabl
spontaneo emissio
de-excite
stimulatec emissio
populatio inversio
optical pu
optical ca
coherent

Albert Einstein
frame of reference
inertial frames of reference
Lorentz transformations
Galilean transformations

the silvered surfaces. This is called *cavity oscillation.*. During n such oscillations, photons multiply n-fold. *Light intensity* becomes very high. Finally these photons get out of the medium through the *semi silvered surface* of the cavity. Emitted light is highly intense, directional and coherent. It is called laser.

Special relativity

It is formulated by Albert Einstein and published in 1905.

Inertial frame of reference : A frame of reference refers to the set consisting of an observer, a coordinate system and a clock. It is used for the observation of physical phenomena and its mathematical description as physical laws. A frame of reference is said to be inertial if newtons laws of motion are valid in them.

The basic postulates of relativity: In all inertial frames of reference

1. The laws of physics have the same form.
2. The velocity of light is a constant.

Lorentz transformations: Let S and S' are two inertial frames of reference with S' moving with velocity v relative to S along $+x$ direction. Then an observer in S will find the coordinates of an event (x, y, z, t) as

$$x' = \gamma(x - vt),\ y' = y,\ z' = z,\ t' = \gamma\left(t - \frac{xv}{c^2}\right) \text{ where } \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

These are Lorentz transformations. Similarly, an observer in S' will find S moving with velocity $-v$ and an event (x', y', z', t') as

$$x = \gamma(x' + vt'),\ y = y',\ z = z',\ t = \gamma\left(t' + \frac{x'v}{c^2}\right)$$

These are inverse Lorentz transformations. When $v << c$, the velocity of light $\gamma \approx 1$ and $xv/c^2 \approx 0$. In this case

$$x' = x - vt,\ y' = y,\ z' = z,\ t' = t$$

These are called Galilean transformations.

Time dilation: Let inertial frame S' be moving with velocity $+v$ along x-axis of inertial frame S'. Suppose two events occur at the same x' at times t'_1 and t'_2 in S'. Let t_1 and t_2 be the time instants observed using two clocks at $x_1 = \gamma(x' + vt'_1)$ and $x_2 = \gamma(x' + vt'_2)$ in S. Then

$$t_1 = \gamma(t'_1 + \frac{x'v}{c^2}),\ t_2 = \gamma(t'_2 + \frac{x'v}{c^2})$$

$$\Delta t = t_2 - t_1 = \gamma(t'_2 - t'_1) = \gamma \Delta t'$$

Since $\gamma > 1$ for $v > 0$, $\Delta t > \Delta t'$. Moving clock which is in S' run slow when observed from S. This is called time dilation. Time interval is minimum in the frame in which the clock is at rest. It is called proper time denoted by $\Delta\tau$.

Length contraction: Consider a rod moving with velocity v along $+x$ of S. Let the x-coordinates x_1 and x_2 of the two ends of rod are measured at the same time t in S. Let x'_1 and x'_2 be the corresponding quantities in S' in which the rod is at rest.

$$x'_1 = \gamma(x_1 - vt),\ x'_2 = \gamma(x_2 - vt)$$

$$\Delta x' = x'_2 - x'_1 = \gamma(x_2 - x_1) = \gamma \Delta x$$

Since $\gamma > 1$ for $v > 0$, $\Delta x' > \Delta x$. Moving rod which is in S' has smaller length when observed from S. This is called length contraction. Length is maximum in the frame in which the rod is at rest. It is called proper length denoted by L_0.

Twin paradox: Consider twins A and B. Let A remains on earth while B travels to a star one light year away with a velocity 0.8c. Using the Lorentz factor $\gamma = 1/\sqrt{1 - 0.64} = 1.6667$ A calculates that B reach there in 1/0.8=1.25years. B and his clock are moving and hence slow. Due to length contraction, the distance to the star is $1/\gamma = 0.6$ *lightyear*. Since velocity of light is constant, the time taken for travel is $0.6/0.8 = 0.75$ *year*. If he immediately turns around and come back to earth with the same speed A measures total time of travel as 2.5 years and B as 1.5 years. So B is younger by one year due to his journey. From principle of equivalence only their relative velocity is measurable. From the point of view of B, it is A that is travelling in the opposite direction with velocity 0.8c. So A must be younger by one year. This is the twin paradox.

Need for experimentation

Experimentation in physics is needed for the following reasons.

- To verify a law deduced theoretically. eg. Curvature of spacetime near the sun.
- To study the behaviour of a physical system under various external conditions. eg. The variation of the state of a gas under different conditions of temperature and pressure.

time dilat
proper tir
length contrac
proper ler
Lorentz fa
twin para
Curvature spacetir

experimental design

- To establish the correlation between various properties of a physical system. eg. The strength of nuclear force and the orientation of nucleon spins.

Design of an experiment

Ronald A. Fisher proposed a methodology for experimental design. It involves the following steps.

1. *Comparison:* If the value of a quantity is too small, it is easier to compare it with an identical but different physical quantity.
2. Replication: Measurements are usually subject to variation and uncertainty. Hence full experiments are often replicated to help identify the sources of variation, to better estimate the true value.
3. Blocking: Blocking refers to dividing a team of researchers into identical groups to do the same experiment. It reduces irrelevant sources of variation for greater precision.
4. Controlled input: The input parameters are so controlled that variation is kept as small as possible. Such controlling mechanism is mandatory for any experiment.
5. Accurate measurement of output: Devices must be set up for the accurate measurement of values of output parameters.
6. Analysis of Output: Methods must be made available to analyse and find correlations, independence between input and output data.

Interaction between physics and technology

All forms of industry are various sections of physics or chemistry applied on a large scale.eg.Dynamos and motors are invented because of Faraday's discovery of electromagnetic induction. It is also true that many concepts of physics are discovered while solving technical problems.eg. The necessity to convert solar energy into electricity led to the development of materials for more efficient photovoltaic solar panels. It is also observed that many fundamental problems lose relevance when they cease to have technical importance. Electrification by friction lost importance when voltaic cells were invented.

4 Matrices

Basic ideas of matrices

Definition 1 (Matrix) *A matrix is an ordered set of real or complex numbers listed in rectangular form as rows and columns.*

eg. If $A = \begin{pmatrix} 1 & 2 & 4 & 9 \\ 2 & 1 & 0 & -1 \\ 0 & 5 & 2 & 4 \end{pmatrix}$, then A is a matrix. A has 12 elements distributed in 3 rows and 4 columns. The dimension of A is written as 3×4 which is read as '*three by four*. A $m \times n$ matrix has m rows and n columns. Its general form can be represented as

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = [a_{ij}], i = 1, 2, .., m, \ j = 1, 2, .., n$$

If $m = n$, it is a *square matrix*, otherwise, a *rectangular matrix*. If $m = 1$ for an $(m \times n)$ matrix,it has only one row and is called *row matrix*. If $n = 1$, it is a *column matrix*. $a_{ij} = 0$ for all i, j, it is called a *null matrix*. In a square matrix, elements a_{ii} are called *diagonal elements*. The sum of diagonal elements of a square matrix $\sum_i a_{ii}$ is called its *trace*. If $a_{ij} = 0$ for all $i \neq j$ and at least one of $a_{ii} \neq 0$, then it is called a *diagonal matrix*. If diagonal elements are all equal in a diagonal matrix, it is called a *scalar matrix*. If diagonal elements of a scalar matrix are unity,($a_{ii} = 1$), it is called a *unit matrix* or *identity matrix*. If the elements of a matrix are complex ($z = x + iy, \ i = \sqrt{-1}, \ x, y \ real$) it is called a *complex matrix*. For any complex variable $z = x + iy$ a complex conjugate $z^* = x - iy$ exists. Hence for any complex matrix $A = [a_{ij}]$ a *conjugate matrix* $A^* = [a^*_{ij}]$ exists.

square mo
trace
diagonal n
symmetri
skew-sym matrix
hermitian

Definition 2 (Equality of matrices) *Two matrices are said to be equal, if their corresponding elements are equal.$A = [a_{ij}]$ and $B = [b_{ij}]$ are equal if $a_{ij} = b_{ij}$ for all i, j.*

Definition 3 (Transpose of a matrix) *If rows and columns of a $(m \times n)$ matrix A are interchanged, it will be a $(n \times m)$ matrix B. If $A = [a_{ij}]$, then $B = [b_{ij}] = [a_{ji}]$. The new matrix B is called transpose of A represented as A^T.*

eg. If $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $A^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$. If matrix $A = A^T$, it is called a *symmetric matrix*.eg. $S = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$, $S^T = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix} = S$ This is possible only for a square matrices. If $A = -A^T$, it is a *skew-symmetric matrix*. For a skew-symmetric matrix, $a_{ij} = -a_{ij}$. Hence diagonal elements $a_{ii} = -a_{ii} = 0$. eg. $S = \begin{pmatrix} 0 & 2 & -3 \\ -2 & 0 & 5 \\ 3 & -5 & 0 \end{pmatrix}$, $S^T = \begin{pmatrix} 0 & -2 & 3 \\ 2 & 0 & -5 \\ -3 & 5 & 0 \end{pmatrix} = -S$. The conjugate transpose of a complex matrix is called its *hermitian conjugate*. $(A^*)^T = (A^T)^* = A^\dagger$. If $A = A^\dagger$, it is called a *hermitian matrix*. eg. If $H = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $H^\dagger = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $H = H^\dagger$. H is hermitian If $A = -A^\dagger$, it *determinant*

is called a *skew-hermitian matrix*.eg.$H = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, $H^\dagger = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = -H$. H is skew-hermitian

Definition 4 (Determinant of a square matrix) *The determinant of a square matrix A is a number associated with A which is a function of the elements of A.*

It is represented as

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

For a 2×2 matrix $A = [a_{ij}]$, it is given by $det.A = a_{11}a_{22} - a_{12}a_{21}$. For a 3×3 matrix

$$|A| = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

The determinant shows the following properties.

1. $\det(I_n) = 1^n = 1$ where I_n is the $(n \times n)$ identity matrix. Hence $\det(cI_n) = c^n 1^n = c^n$ for a number c, real or complex.
2. $\det(A^T) = \det(A)$
3. For conformable square matrices A and B, $\det(AB) = \det(A)\det(B)$
4. If A^{-1} is the inverse of A,

$$AA^{-1} = I, \det(AA^{-1}) = 1,\ \det(A^{-1}) = \frac{1}{\det(A)}$$

5. If c is a number, for an $(n \times n)$ matrix A, $\det(cA) = c^n \det(A)$
6. If A is a triangular matrix($a_{i,j} = 0$ for $i > j$ or $i < j$), then its determinant equals the product of the diagonal elements

$$\det(A) = a_{11}a_{22}\cdots a_{nn} = \prod_{i=1}^{n} a_{ii}$$

Definition 5 (Sum of matrices) *The sum of a set of given matrices is a matrix whose elements are the sum of corresponding elements of matrices of the set.*

This requires all matrices of the set must have the same dimension. Such matrices are said to be conformable for addition. If $D = A + B + C$, then $d_{ij} = a_{ij} + b_{ij} + c_{ij}$. The law of addition of matrices leads to the properties.

1. Associativity $(A + B) + C = A + (B + C)$

2. Existance of null matrix. $A + 0 = A = 0 + A$

3. Existance of additive inverse (opposite matrix) $A + (-A) = 0$

4. Commutativity $A + B = B + A$

Definition 6 (Scalar multiplication of matrices) *A scalar is a mathematical object which can be expressed using a single number. (eg. mass). If $A = [a_{ij}]$ and α a scalar, then $\alpha A = [\alpha a_{ij}]$.*

This law of scalar multiplication leads to the following properties.

$(1)(\alpha+\beta)A = \alpha A+\beta A(2)(\alpha\beta)A = \alpha(\beta A) = \beta(\alpha A)(3)\alpha(A+B) = (\alpha A+\alpha B(4)(\alpha A)^T = \alpha A^T$

Definition 7 (Product of matrices) *If $A = [a_{ij}], B = [b_{ij}]$ are two matrices, then their matrix product is defined as a matrix $C = [c_{ik}]$ where $c_{ik} = \sum_j a_{ij}b_{jk}$*

Since summation occurs over column subscript of A and row subscript of B, number of columns of A and number of rows of B must be equal. Such matrices are said to *conformable for matrix multiplication.* A $(m\times n)$ and a $(p\times q)$ matrices are conformable for matrix multiplication, if $n = p$ and the matrix product is a $(m \times q)$ matrix. For example

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{13}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}$$

The product may be written as

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{13}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{3} a_{1j}b_{j1} & \sum_{j=1}^{3} a_{1j}b_{j2} \\ \sum_{j=1}^{3} a_{2j}b_{j1} & \sum_{j=1}^{3} a_{3j}b_{j3} \end{pmatrix}$$

$C = AB$ can be written as $[c_{ik}] = [\sum_j a_{ij}b_{jk}]$ If $AB = C$ is a null matrix, B is called the *zero divisor* of A. Matrix multiplication has the following properties.

1. In general $AB \neq BA$(Non-commutativity)

2. $A(BC) = (AB)C$ (Associativity)

3. A(B+C)=AB+AC (Distributivity)

Definition 8 (Unitary matrices) *A square matrix A is unitary if $AA^\dagger = A^\dagger A = I$*

Hence $A^{-1} = A^\dagger$. Hermitian conjugate is also the inverse.
eg. If $U = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $U^\dagger = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $UU^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Representation of vectors as column matrix

A vector $\vec{r}$ in 3-dimensions can be written as the vector sum $\vec{r} = \hat{\imath}x + \hat{\jmath}y + \hat{k}z$ where $(\hat{\imath}, \hat{\jmath}, \hat{k})$ are unit vectors along (x, y, z)-axes. It may be represented as an ordered array(column matrix)

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

If x, y, z are real,

$$(\vec{r})^T \vec{r} = (x, y, z) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2$$

which is square of length of $\vec{r}$. If x, y, z are complex,

$$(\vec{r})^\dagger \vec{r} = (x^*, y^*, z^*) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = |x|^2 + |y|^2 + |z|^2$$

which is also the square of length of $\vec{r}$

Cramer's rule

Given a set of three linear equations

$$a_1x + b_1y + c_1z = d_1; a_2x + b_2y + c_2z = d_2; a_3x + b_3y + c_3z = d_3$$

which may be written in the matrix form,

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

Consider the four determinants

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, D_x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}, D_y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}, D_z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}$$

Then the solution for (x, y, z) is given by $x = D_x/D, y = D_y/D, z = D_z/D$. This can be extended for any number of variables x. This is called Cramer's rule.
eg. Let the equations to be solved are $2x + 3y - z = 5,\ 3x - 2y + z = 2,\ x + y + z = 6$.
The coefficient matrix, $A = \begin{pmatrix} 2 & 3 & -1 \\ 3 & -2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. The determinants

$$D = \begin{vmatrix} 2 & 3 & -1 \\ 3 & -2 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -17,\ D_x = \begin{vmatrix} 5 & 3 & -1 \\ 2 & -2 & 1 \\ 6 & 1 & 1 \end{vmatrix} = -17$$

$$D_y = \begin{vmatrix} 2 & 5 & -1 \\ 3 & 2 & 1 \\ 1 & 6 & 1 \end{vmatrix} = -34, \; D_z = \begin{vmatrix} 2 & 3 & 5 \\ 3 & -2 & 2 \\ 1 & 1 & 6 \end{vmatrix} = -51$$

$x = D_x/D = (-17/-17) = 1, y = D_y/D = (-34/-17) = 2, z = D_z/D = (-51/-17) = 3$

Problem: Solve the following using Cramer's Rule.

$$2x + 3y + 4z = 2, \; 5x + 6y + 5z = 0, \; 2.5x + 0.1y + 7.5z = 5.0$$

Eigen Values and Eigen Vectors

eigenvalu
eigenvecto
degenerat
eigenvalu spectru
characteri equatio
polynomia equatio

A matrix is an operator. When it operates on a column vector, it creates a new vector. $A\vec{X} = \vec{Y}$. For example, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ There may exist certain non-zero vectors for which $A\vec{X} = \lambda\vec{X}$ where λ is a constant number. λ is called an *eigenvalue* or characteristic value of the matrix A. $\vec{X}$ is called an *eigenvector* corresponding to the eigenvalue λ. If more than one eigenvector correspond to the same eigenvalue, the eigenvalue is said to be degenerate. The set of eigenvalues of a matrix (an operator) is called its *eigenvalue spectrum.*

Determination of eigenvalues and eigenvectors of a matrix: There are two parts to every eigenvalue problem. First, we compute the eigenvalue λ of the given the matrix A. Then, we compute an eigenvector X for each eigenvalue λ. Let I be a unit matrix of the same dimension as A.

$$A\vec{X} = \lambda\vec{X} = \lambda I\vec{X}, \; (A - \lambda I)\vec{X} = 0$$

Since X is a non-zero vector $(A - \lambda I) = 0$. If A is a $n \times n$ matrix,

$$(A - \lambda I) = \begin{pmatrix} a_{11} - \lambda & a_{12} - 0 & ... & a_{1n} - 0 \\ a_{21} - 0 & a_{22} - \lambda & ... & a_{2n} - 0 \\ ... & ... & ... & ... \\ a_{n1} - 0 & a_{n2} - 0 & ... & a_{nn} - \lambda \end{pmatrix} = \textit{null matrix}$$

$$\det{(A - \lambda I)} = \begin{vmatrix} a_{11} - \lambda & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} - \lambda & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} - \lambda \end{vmatrix} = 0$$

This equation is called the characteristic equation of matrix A. The expansion of the determinant gives an n^{th} order polynomial equation in λ. It may be written as

$$c_n\lambda^n + c_{n-1}\lambda^{n-1} + ... + c_1\lambda^1 + c_0 = 0$$

normalisation condition

where c_i are functions of elements of A. This polynomial has n roots which are values of λ. It can be shown that, the eigenvalues satisfy the following relations.
(1)The sum of the eigenvalues equals the trace of the matrix.

$$\lambda_1 + \lambda_2 + \lambda_3 + ... + \lambda_n = a_{11} + a_{22} + a_{33} + ... + a_{nn}$$

(2) The product of the eigenvalues equals the determinant of the matrix.

$$\lambda_1\lambda_2\lambda_3....\lambda_n = \det A$$

Once the eigenvalues have been found, corresponding eigenvectors can be found from the system $A\vec{X} = \Lambda\vec{X}$. Since the system is homogeneous, if X is an eigenvector of A, then kX, where k is any constant (not zero), is also an eigenvector of A corresponding to the same eigenvalue. For example, to find the eigenvalues and eigenvectors of the real symmetric matrix $A = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 1 & -3 \\ 3 & -3 & -3 \end{pmatrix}$ we solve the determinant equation $\begin{pmatrix} 1-\lambda & 1 & 3 \\ 1 & 1-\lambda & -3 \\ 3 & -3 & -3-\lambda \end{pmatrix}$ Expanding out this determinant gives

$$(1-\lambda)[(1-\lambda)(-3-\lambda)-(-3)(-3)]-1[1(-3-\lambda)-(-3)(3)]+3[1(-3)-(1-\lambda)(3)] = 0$$

which simplifies to give

$$(1-\lambda)(\lambda^2 + 2\lambda - 12) + (\lambda - 6) + 3(3\lambda - 6) = 0$$

$$(\lambda - 2)(\lambda - 3)(\lambda + 6) = 0$$

Hence the roots of the characteristic equation, which are the eigenvalues of A, are $\lambda_1 = 2, \lambda_2 = 3, \lambda_3 = -6$. We note that, as expected, For the first root, $\lambda_1 = 2$, a suitable eigenvector X_1 , with elements x_1, x_2, x_3 must satisfy $AX^1 = 2X^1$ That is, $x_1 + x_2 + 3x_3 = 2x_1,\ x_1 + x_2 - 3x_3 = 2x_2,\ 3x_1 - 3x_2 - 3x_3 = 2x_3$ These three equations are consistent and give $x_3 = 0, x_1 = x_2 = k$, where k is any non-zero number. A suitable eigenvector would thus be $X^1 = \begin{pmatrix} k \\ k \\ 0 \end{pmatrix}$. If we apply the normalisation condition, we require $(X^1)^T X^1 = k^2 + k^2 + 0^2 = 1 \ or\ k = \sqrt{1/2}$. Hence $X_1 = \begin{pmatrix} \sqrt{1/2} \\ \sqrt{1/2} \\ 0 \end{pmatrix} = \sqrt{1/2}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ Repeating for $\lambda_2 = 3$ and $\lambda_3 = -6$, two normalised eigenvectors $X^2 = \sqrt{1/3}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $X^3 = \sqrt{1/6}\begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$ will be obtained.

Problem: Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ hermitian
unitary m

Note:
(1)The eigenvalues of a hermitian matrix are real and eigenfunctions belonging to different eigenvalues are orthogonal.
(2)The eigenvalues of a unitary matrix are $\lambda = e^{i\alpha}$ so that $|\lambda| = \sqrt{\lambda.\lambda^*} = \sqrt{e^{i\alpha}.e^{-i\alpha}} = \pm 1$.

point function
scalar point functions
vector point functions
scalar point function
vector field

5 Vector calculus

Definition 9 (Point function) *A point function $v = f(P)$ is a function that assigns some number or value v to each point P of some region R of space. Point functions are classified into scalar point functions and vector point functions.*

Definition 10 (Scalar point function) *A scalar point function is a function that assigns a real number (a scalar) to each point of some region of space. If to each point (x, y, z) of a region R in space there is assigned a real number $v = \phi(x, y, z)$, then ϕ is called a scalar point function.*

Examples. 1. The temperature distribution within some body at any instant of time. 2. The density distribution within some fluid at any instant of time.

Definition 11 (Scalar field) *A scalar point function defined over some region is called a scalar field.*

A scalar field that varies with time would have the representation $v = \phi(x, y, z, t)$. A scalar field which is independent of time is called a stationary or steady-state scalar field.

Definition 12 (Vector point function) *A vector point function is a function that assigns a vector to each point of some region of space.*

If to each point (x, y, z) of a region R in space, there is assigned a vector $\vec{F} = \vec{F}(x, y, z)$, $\vec{F}$ is called a vector point function. Such a function would have a representation $\vec{F} = \hat{i}u(x, y, z) + \hat{j}v(x, y, z) + \hat{k}w(x, y, z)$

Definition 13 (Vector field) *A vector point function defined over some region is called a vector field.*

A vector field which is independent of time is called a stationary or steady-state vector field. Examples. 1. Gravitational field of the earth. 2. Electric field about a current-carrying wire. 3. Magnetic field generated by a magnet. 3. Velocity at different points within a moving fluid.

Definition 14 (Partial derivative) *Partial derivative of a function $f(x_1, x_2, ..., x_n)$ of several independent variables x with respect to x_i is defined as the derivative when all variables x except x_i are held fixed during differentiation.*

$$\frac{\partial f(x_1, x_2, ..., x_n)}{\partial x_i} = \lim_{\delta x_i \to 0} \frac{f(x_1, x_2, ..., x_i + \delta x_i, ..., x_n) - f(x_1, x_2, .., x_i, .., x_n)}{\delta x_i}$$

For example, if $f(x, y, z) = x^2 + xy + yz + zx + z^2$

$$\frac{\partial f}{\partial x} = 2x + y + z, \ \frac{\partial f}{\partial y} = x + z, \ \frac{\partial f}{\partial z} = y + x + 2z$$

Definition 15 (Gradient) *The gradient is a vector operator acting on scalar point functions giving a vector function.*

vector ope
gravitatio
potenti
potential
outward n

In Cartesian coordinates, it has the form $\nabla = \hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}$ where $(\hat{i}, \hat{j}, \hat{k})$ are unit vectors along x, y, z axes respectively. if $f(x, y, z)$ is a scalar point function,

$$\nabla f = \hat{i}\frac{\partial f}{\partial x} + \hat{j}\frac{\partial f}{\partial y} + \hat{k}\frac{\partial f}{\partial z}$$

The magnitude of ∇f at any point (x_0, y_0, z_0) is

$$|\nabla f|_{(x_0,y_0,z_0)} = \left[\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2}\right]_{(x_0,y_0,z_0)}$$

The direction in which $|\nabla f|$ is maximum is the direction of ∇f.

Example of gradient: In general, the gravitational potential V at any point $\vec{r} = (x, y, z)$ due to a stationary mass is a function of (x, y, z). Hence any change in V is a consequence of the individual changes in x, y and z. The change in V due to change dx in x is given by the product of rate of change of V relative to x and change in x. That is $(dV)_x = \frac{\partial V}{\partial x}dx$. Similarly,$(dV)_y = \frac{\partial V}{\partial y}dy$ and $(dV)_z = \frac{\partial V}{\partial z}dz$. Hence,

$$dV = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz = \left(\hat{i}\frac{\partial V}{\partial x} + \hat{j}\frac{\partial V}{\partial y} + \hat{k}\frac{\partial V}{\partial z}\right).(\hat{i}dx + \hat{j}dy + \hat{k}dz) = \nabla V.d\vec{r}$$

dV may be calculated using the idea of work done. The potential energy of a body decreases if it moves in the direction of field of force. Gravitational potential V is potential energy per unit mass. dV is negative if gravitational force $\vec{F}$ and infinitesimal displacement $d\vec{r}$ are parallel. Hence $dV = -\vec{F}.d\vec{r}$. Comparing expressions for dV, one gets $-\vec{F} = \nabla V$ or $\vec{F} = -\nabla V$ where $F_x = \frac{\partial V}{\partial x}, F_y = \frac{\partial V}{\partial y}, F_z = \frac{\partial V}{\partial z}$.

The direction of ∇V: Let $V \equiv V(x, y, z)$ is a scalar point function. $V(x, y, z) = C$ is a surface S if C is a constant. Then a change in V due to a small change $d\vec{r} = \hat{i}x + \hat{j}y + \hat{k}z$ tangential to the surface S is given by $dV = \nabla V.d\vec{r} = 0$. This means ∇V and $d\vec{r}$ are normal to each other. Since $d\vec{r}$ is tangential to S, ∇V is normal to S. Consider two surfaces $V(x, y, z) = C$ and $V(x, y, z) = C + \epsilon$ where ϵ is very small. Let $d\vec{r_1}$ be a small vector from $V = C$ to $V = C + \epsilon$. Change of V between the surfaces

$$dV = C + \epsilon - C = \epsilon = \nabla V.d\vec{r_1}$$

. ϵ is maximum when ∇V and $d\vec{r_1}$ are parallel. Hence ∇V is along the maximum change of V which is the outward normal to the surface $V = Constant$.

line integral
surface integral

Definition 16 (Line integral) *A line integral is an integral where the function f to be integrated is evaluated along a curve C. If f is a scalar point function, the line integral is given by*

$$\int_C f(\vec{r})d\vec{r} = \lim_{N\to\infty}\sum_{i=1}^{N} f(\vec{r}_i)\Delta\vec{r}_i$$

where $\Delta\vec{r}_i$ is a line element along C. If f is a vector point function, two types of line integrals are defined.

$$\int_C \vec{f}(\vec{r}).d\vec{r} = \lim_{N\to\infty}\sum_{i=1}^{N} \vec{f}(\vec{r}_i).\Delta\vec{r}_i$$

$$\int_C \vec{f}(\vec{r})\times d\vec{r} = \lim_{N\to\infty}\sum_{i=1}^{N} \vec{f}(\vec{r}_i)\times\Delta\vec{r}_i$$

Problem: Integrate $f(x,y) = x^2 + xy$ from $(0,1)$ to $(1,3)$ along $y = 2x^2 + 1$.
The function $f(x,y)$ along $y = 2x^2 + 1$ has the form $f(x,y) = x^2 + x(2x^2+1) = 2x^3 + x^2 + x$.

$$\int_0^1 (2x^3 + x^2 + x)dx = \left[\frac{x^4}{2} + \frac{x^3}{3} + \frac{x^2}{2}\right]_0^1 = \frac{1}{2} + \frac{1}{3} + \frac{1}{2} = \frac{4}{3}$$

If the C is a closed curve, the line integral is represented as $\oint_C$

Definition 17 (Surface integral) *A surface integral is an integral where the function f to be integrated is evaluated on a surface S. If f is a scalar point function, the surface integral is given by*

$$\int_C f(\vec{r})d\vec{s} = \lim_{N\to\infty}\sum_{i=1}^{N} f(\vec{r}_i)\Delta\vec{s}_i$$

where $\Delta\vec{s}_i$ is an elementary area on S. If f is a vector point function, two types of surface integrals are defined.

$$\int_C \vec{f}(\vec{r}).d\vec{s} = \lim_{N\to\infty}\sum_{i=1}^{N} \vec{f}(\vec{r}_i).\Delta\vec{s}_i$$

$$\int_C \vec{f}(\vec{r})\times d\vec{s} = \lim_{N\to\infty}\sum_{i=1}^{N} \vec{f}(\vec{r}_i)\times\Delta\vec{s}_i$$

If the S is a closed surface, the surface integral is represented as $\oint_S$

Definition 18 (Volume integral) *A volume integral is an integral where the function f to be integrated is evaluated within a volume τ. If f is a scalar point function, the volume integral is a scalar. It is given by*

$$\int_C f(\vec{r})d\tau = \lim_{N\to\infty}\sum_{i=1}^{N} f(\vec{r}_i)\Delta\tau_i$$

where $\Delta\tau_i$ is an elementary volume in τ. If f is a vector point function, the volume integral is a vector.

$$\int_C \vec{f}(\vec{r})d\tau = \lim_{N\to\infty}\sum_{i=1}^{N} \vec{f}(\vec{r}_i)\Delta\tau_i$$

volume in
divergenc

Definition 19 (Divergence) *The divergence of a vector point function $\vec{f}(\vec{r})$ is defined as*

$$div\vec{f} = \lim_{\delta\tau\to 0}\frac{\oint_S \vec{f}(\vec{r}).d\vec{s}}{\delta\tau}$$

where S encloses τ. Consider a volume bounded by $(\delta x, \delta y, \delta z)$. The surface area $\delta y\delta z$ is normal to x-axis. Let

$$\vec{f}(\vec{r}) = \vec{f}(x,y,z) = \hat{\imath}f_x(x,y,z) + \hat{\jmath}f_y(x,y,z) + \hat{k}f_z(x,y,z)$$

$$\left(\oint_S \vec{f}.d\vec{s}\right)_x = \lim_{\delta x\to 0}[f_x(x+\delta x,y,z) - f_x(x,y,z)]\delta y\delta z$$

Multiplying and dividing by δx

$$\left(\oint_S \vec{f}.d\vec{s}\right)_x = \lim_{\delta x\to 0}\frac{f_x(x+\delta x,y,z) - f_x(x,y,z)}{\delta x}\delta x\delta y\delta z = \frac{\partial f_x}{\partial x}\delta\tau$$

Similarly,

$$\left(\oint_S \vec{f}.d\vec{s}\right)_y = \frac{\partial f_y}{\partial y}\delta\tau,\ \left(\oint_S \vec{f}.d\vec{s}\right)_z = \frac{\partial f_z}{\partial z}\delta\tau$$

$$div\vec{f} = \lim_{\delta\tau\to 0}\frac{1}{\delta\tau}\left[\frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}\right]\delta\tau = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z} = \nabla.\vec{f}$$

Definition 20 (Curl) *The curl of a vector point function in the direction $\hat{n}$ is defined as*

$$curl\ \vec{f}.\hat{n} = \lim_{\delta s\to 0}\frac{\oint_C \vec{f}(\vec{r}).d\vec{r}}{\delta s}$$

where curve C encloses area δs. Let C be a small rectangular element bounded by sides of length δx and δy. The line integral $\oint_C \vec{f}.d\vec{r}$ is given by

$$\oint_C \vec{f}.d\vec{r} = \int_x^{x+\delta x} f_x(x,y,z)dx + \int_y^{y+\delta y} f_y(x+\delta x,y,z)dy + \int_{x+\delta x}^{x} f_x(x,y+\delta y,z)dx + \int_{y+\delta y}^{y} f_y(x,y,z)dy$$

conservative field
Lorentz force
non-conservative force
solenoidal
irrotational

Rearranging

$$\oint_C \vec{f}.d\vec{r} = \int_x^{x+\delta x} f_x(x,y,z)dx - \int_x^{x+\delta x} f_x(x,y+\delta y,z)dx$$
$$+ \int_y^{y+\delta y} f_y(x+\delta x,y,z)dy - \int_y^{y+\delta y} f_y(x,y,z)dy$$

But if δy is small

$$[f_x(x,y,z) - f_x(x,y+\delta y,z)]dx = \frac{[f_x(x,y,z) - f_x(x,y+dy,z)]}{dy}dxdy = -\frac{\partial f_x}{\partial y}dxdy$$

Similarly the other two terms

$$f_y(x+\delta x,y,z)dy - f_y(x,y,z)dy = \frac{\partial f_y}{\partial x}dxdy$$

If $dx\ dy = ds$

$$\oint_C \vec{f}.d\vec{r} = \int_S \left[\frac{\partial f_y}{\partial x} - \frac{\partial f_x}{\partial y}\right] ds = \int_S (\nabla \times \vec{f})_z ds = \int_S (\nabla \times \vec{f}).\hat{k}\ ds$$

$$curl\ \vec{f}.\hat{n} = \lim_{\delta s \to 0} \frac{\oint_C \vec{f}(\vec{r}).d\vec{r}}{\delta s} = (\nabla \times \vec{f}).\hat{k}$$

where $\hat{k}$ is a unit vector along z axis. That is, if the contour C is in the $x-y$ plane, then curl $\vec{f}$ is in the z-direction. Hence $\boxed{curl\ \vec{f} = \nabla \times \vec{f}}$

Theorems of vector integrals

Theorem 1 (Path independence) *The necessary and sufficient condition that $\int_c \vec{f}.d\vec{r}$ is independent of path is that there exists a scalar point function V such that $\vec{f} = \nabla V$.*

Theorem 2 (Path independence) *The necessary and sufficient condition that $\int_c \vec{f}.d\vec{r}$ is independent of path is that $\nabla \times \vec{f} = 0$.*

Such a vector field is called a conservative field. Gravitational field , electrostatic field etc are examples of conservative fields. Frictional force, magnetic Lorentz force etc are non-conservative force fields.

Theorem 3 (Divergence theorem) *If $\vec{f}$ is a continuously differentiable vector field in a region of volume τ bounded by a surface S, then $\oint_s \vec{f}.d\vec{s} = \int_\tau \nabla.\vec{f}d\tau$.*

If $\nabla.\vec{f} = 0$, then $\vec{f}$ is called solenoidal. Magnetic field $\vec{B}$ is solenoidal.

Theorem 4 (Stokes theorem) *If a closed curve C encloses an area S where a vector point function $\vec{f}$ has continuous partial derivatives, then $\oint_C \vec{f}.d\vec{r} = \int_S \nabla \times \vec{f}.d\vec{s}$.*

If $\nabla \times \vec{f} = 0$, then $\vec{f}$ is called irrotational. Conservative fields like electrostatic and gravitational fields are irrotational

Physical interpretation of gradient, divergence and curl

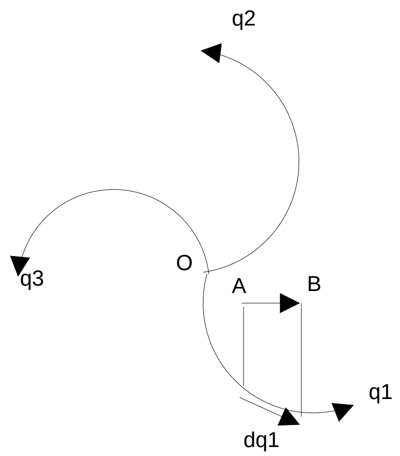
Gradient represents steepness. For example, if temperature distribution in a room is the scalar point function $f(\vec{r}) = T(\vec{r})$, then $\nabla T(\vec{r})$ determines temperature difference directed towards the greatest change in the temperature field . If temperature in the room is same at every point, then gradient will be zero.

In a moving fluid, the velocity at different points form a vector field $\vec{f}(\vec{r}) = \vec{v}(\vec{r})$. Curl and divergence of $\vec{v}(\vec{r})$ are used to express notions of rotation and expansion/compression of a fluid, respectively. Divergence represents flux per unit volume. Suppose there is a liquid flowing through a uniform pipe. Its divergence is zero. If it is coming out and expands then its divergence is positive. If the liquid passes through a funnel and comes out through its tail, then divergence is negative within the funnel. If it is a gas flow, divergence expresses ideas of expansion and compression of the gas. So, if $\nabla.\vec{v}(\vec{r}) > 0$ the tendency is for fluid to move away from the point $\vec{r}$. If $\nabla.\vec{v}(\vec{r}) < 0$ then the fluid is tending to move towards that point (compression). If $\nabla.\vec{v}(\vec{r}) = 0$ then the fluid flow is streamlined.

The direction of the curl of a vector field is the axis of rotation, as determined by the right-hand screw rule. The magnitude of the curl is the magnitude of rotation.

Orthogonal curvilinear coordinate

A set of coordinates $q_1 = q_1(x, y, z),\ q_2 = q_2(x, y, z),\ q_3 = q_3(x, y, z)$ where the directions at any point indicated by q_1, q_2 and q_3 are orthogonal (perpendicular) to each other is referred to as a set of orthogonal curvilinear coordinates.



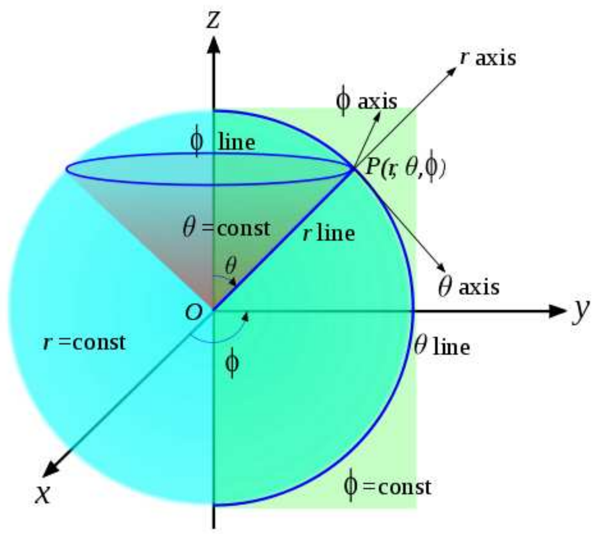
In general, $q_1 = constant,\ q_2 = constant$ and $q_3 = constant$ are curved surfaces. Since axes are intersections of these surfaces, they are curved. Hence associated with each such coordinate, a scale factor must be defined to get a measure of how a change in the coordinate changes the position of a point. If (h_1, h_2, h_3) are the scale factors associated with (q_1, q_2, q_3), then the infinitesimal distance between any pair of points (q_1, q_2, q_3)

and $(q_1 + dq_1, q_2 + dq_2, q_3 + dq_3)$ is given by

$$ds^2 = h_1^2\, dq_1^2 + h_2^2\, dq_2^2 + h_3^2\, dq_3^2$$

Generally scale factors are functions of (q_1, q_2, q_3). Cylindrical polar coordinates and spherical polar coordinates are two systems of orthogonal curvilinear coordinates.

Spherical polar coordinate



The three coordinates are (r, θ, ϕ). Their ranges are $r \geq 0,\ 0 \leq \theta \leq \pi,\ 0 \leq \phi \leq 2\pi$. Elementary length is given by

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2$$

. Therefore, $h_1 = 1, h_2 = r, h_3 = r\sin\theta$. The three surfaces are $r = constant$(spherical), $\theta = constant$(conical) and $\phi = constant$(semicircular).

Cylindrical polar coordinate

The three coordinates are (r, ϕ, z). Their ranges are $r \geq 0,\ 0 \leq \phi \leq 2\pi,\ -\infty \leq z \leq +\infty$. The elementary length is given by

$$ds^2 = dr^2 + r^2 d\phi^2 + z^2,\ h_1 = 1, h_2 = r, h_3 = 1$$

The surfaces $r = constant$ is cylindrical, $\phi = constant$ is plane and $z = constant$ is circular.

Index