Contents 47

Mathematics and Physics Miscellany

Dimensional Analysis in Engineering

47.1

Introduction

Very often, the calculations carried out in engineering and science lead to a result which is not just a pure number but rather a number with a physical meaning, for example so many metres per second or so many kilograms per cubic metre or so many litres. If the numbers which are the input to the calculation also have a physical meaning then we can bring into play a form of analysis (called dimensional analysis) which supplements the arithmetical part of the calculation with an extra analysis of the physical units associated with the numerical values. In its simplest form dimensional analysis just helps in checking that a formula makes sense. For example, consider the *constant acceleration* equation

$$s = ut + \frac{1}{2}at^2$$

On the left-hand side s represents a distance in (say) *metres*. On the right-hand side a being acceleration is so many *metres per second per second*, which, when multiplied by t^2 which is *seconds squared* gives *metres*, as does the other term ut. Thus if somebody copied down the equation wrongly and wrote $\frac{1}{2}at$ (instead of $\frac{1}{2}at^2$) then we could surmise that there must be a mistake. This simple example illustrates the checking of an equation for dimensional homogeneity. However, it turns out that this passive **checking** aspect of dimensional analysis can be enhanced to give a useful **deductive** aspect to the method, so that it is possible to go some way towards finding out useful information about the equations which describe a system even when they are not fully known in advance. The following pages give many examples and exercises which will show how far it is possible to progress by adding an analysis of physical units to the usual arithmetic of numbers.

Prerequisites

Before starting this Section you should . . .

- be fully familiar with the laws of indices and with negative and fractional indices
- be able to solve sets of linear equations involving up to three variables

Learning Outcomes

On completion you should be able to . . .

- check the dimensional validity of a given equation
- determine which combinations of physical variables are likely to be important in the equations which describe a system's behaviour

1. Introduction to dimensional analysis

It is well known that it does not make sense to **add** together a mass of, say 5 kg and a time of 10 seconds. The same applies to **subtraction**. Mass, length and time are different from one another in a fundamental way but it is possible to **divide** length by time to get a **velocity**. It is also possible to **multiply** quantities, for example length multiplied by length gives **area**.

We note that, strictly speaking, we multiply a **number** of metres by a **number** of metres to get a **number** of square metres; however the terminology 'length times length equals area' is common in texts dealing with dimensional analysis.

We use the symbol [ ] to indicate the **dimensions** of a quantity:

$$[\text{mass}] = \mathsf{M} \qquad [\text{length}] = \mathsf{L} \qquad [\text{time}] = \mathsf{T}$$

The dimensions of other quantities can be written in terms of M, L and T using the definition of the particular quantity. For example:

$$[\text{area}] = [\text{length} \times \text{length}] = \mathsf{L}^2$$

$$[\text{density}] = \left[\frac{\text{mass}}{\text{volume}}\right] = \frac{\mathsf{M}}{\mathsf{L}^3} = \mathsf{ML}^{-3}$$

Some 'rules' can help in calculating dimensions:

Rule 1. Constant numbers in formulae do not have a dimension and are ignored.

Rule 2. Angles in formulae do not have a dimension and are ignored.

Newton's second law relates force, mass and acceleration, and we have:

$$[\text{force}] = [\text{mass} \times \text{acceleration}] = [\text{mass}] \times [\text{acceleration}] = \mathsf{MLT}^{-2}$$

$$[\text{work}] = [\text{force} \times \text{distance}] = \mathsf{MLT}^{-2} \times \mathsf{L} = \mathsf{ML}^2\mathsf{T}^{-2}$$

$$[\text{kinetic energy}] = \left[\frac{1}{2}\,\text{mass} \times (\text{velocity})^2\right] = \mathsf{M}(\mathsf{LT}^{-1})^2 = \mathsf{ML}^2\mathsf{T}^{-2}$$

Note that the sum and product rules for indices have been used here and that the dimensions of 'work' and 'kinetic energy' are the same.

Rule 3. In formulae only physical quantities with the same dimensions may be added or subtracted.

That is the end of the theory - the rules. Next is a list of the dimensions of useful quantities in mechanics. All come from the definition of the relevant quantity.

$[\text{mass}] = \text{M}$
$[\text{length}] = \text{L}$
$[\text{time}] = \text{T}$
$[\text{area}] = [\text{length} \times \text{length}] = \text{L}^2$
$[\text{volume}] = [\text{area} \times \text{length}] = \text{L}^3$
$[\text{density}] = [\text{mass/volume}] = \text{M/L}^3 = \text{ML}^{-3}$
$[\text{velocity}] = [\text{length/time}] = \text{L/T} = \text{LT}^{-1}$
$[\text{acceleration}] = [\text{velocity/time}] = \text{LT}^{-1}/\text{T} = \text{LT}^{-2}$
$[\text{force}] = [\text{mass} \times \text{acceleration}] = \text{MLT}^{-2}$
$[\text{moment of force, torque}] = \text{MLT}^{-2} \times \text{L} = \text{ML}^2\text{T}^{-2}$
$[\text{impulse}] = [\text{force} \times \text{time}] = \text{MLT}^{-2} \times \text{T} = \text{MLT}^{-1}$
$[\text{momentum}] = [\text{mass} \times \text{velocity}] = \text{MLT}^{-1}$
$[\text{work}] = [\text{force} \times \text{distance}] = \text{ML}^2\text{T}^{-2}$
$[\text{kinetic energy}] = [\text{mass} \times (\text{velocity})^2] = \text{ML}^2\text{T}^{-2}$
$[\text{power}] = [\text{work/time}] = \text{ML}^2\text{T}^{-3}$

Task

State the dimensions, in terms of M, L and T of the quantities:

(a) pressure (force per unit area)
(b) line density (mass per unit length)
(c) surface density (mass per unit area)
(d) frequency (number per unit time)
(e) angular velocity (angle per unit time)
(f) angular acceleration (rate of change of angular velocity)
(g) rate of loss of mass (mass per unit time)

Your solution

Answer
(a) $\text{ML}^{-1}\text{T}^{-2}$ (b) ML^{-1} (c) ML^{-2} (d) T^{-1} (e) T^{-1} (f) T^{-2} (g) MT^{-1}

2. Simple applications of dimensional analysis

Application 1: Checking a formula

A piano wire has mass m, length l, and tension F. Which of the following formulae for the period of vibration, t, could be correct?

$$(a) \quad t = 2\pi\sqrt{\frac{ml}{F}} \qquad \text{or} \qquad (b) \quad t = 2\pi\sqrt{\frac{F}{ml}}$$

The dimensions of the expressions on the right-hand side of each formula are studied:

(a)
$$\begin{aligned}\left[2\pi\sqrt{\frac{ml}{F}}\right] &= [2\pi]\left[\frac{ml}{F}\right]^{\frac{1}{2}} \\ &= \left(\frac{\mathsf{ML}}{\mathsf{MLT}^{-2}}\right)^{\frac{1}{2}} \\ &= \left(\frac{1}{\mathsf{T}^{-2}}\right)^{\frac{1}{2}} \\ &= (\mathsf{T}^2)^{\frac{1}{2}} \\ &= \mathsf{T}\end{aligned}$$

which **is consistent** with the dimensions of t, a period of vibration.

(b)
$$\begin{aligned}\left[2\pi\sqrt{\frac{F}{ml}}\right] &= [2\pi]\left[\frac{F}{ml}\right]^{\frac{1}{2}} \\ &= \left(\frac{\mathsf{MLT}^{-2}}{\mathsf{ML}}\right)^{\frac{1}{2}} \\ &= \left(\frac{\mathsf{T}^{-2}}{1}\right)^{\frac{1}{2}} \\ &= \mathsf{T}^{-1}\end{aligned}$$

which is **not consistent** with the dimensions of t, a period of vibration.

We can conclude that the second formula cannot be correct, whereas the first formula could be correct. (Actually, for case (b) we can see that the reciprocal of the studied quantity **would** be a time).

Task

Use the method of dimensions to check whether the following formulae are dimensionally valid where u and v are velocities, g is an acceleration, s is a distance and t is a time.

(a) $v^2 = u^2 + 2gs$

(b) $v = u - gt$

(c) $s = 0.5gt^2 + ut$

(d) $t = \sqrt{\dfrac{2s}{g}}$

Your solution

Answer

All are dimensionally valid.

Task

When a tension force T is applied to a spring the extension x is related to T by Hooke's law, $T = kx$, where k is called the stiffness of the spring. What is the dimension of k? Verify the dimensional correctness of the formula $t = 2\pi\sqrt{\dfrac{m}{k}}$ for the period of oscillation of a mass m suspended by a spring of stiffness k.

Your solution

Answer

Since T is a force, it has dimensions of $[\mathsf{M}][\mathsf{L}][\mathsf{T}]^{-2}$.

So $k = T/x$ has dimensions $[\mathsf{M}][\mathsf{L}][\mathsf{T}^{-2}]/[\mathsf{L}] = [\mathsf{M}][\mathsf{T}]^{-2}$.

So $\sqrt{m/k}$ has dimensions of $[\mathsf{M}]^{1/2}/([\mathsf{M}]^{1/2}[\mathsf{T}]^{-1}) = [\mathsf{T}]$

Application 2: Deriving a formula

A simple pendulum has a mass m attached to a string of length l and is at a place where the acceleration due to gravity is g. We suppose that the periodic time of swing t of the pendulum depends jointly on $m,\ l$ and g and the aim is to find a formula for t in terms of $m,\ l$ and g which is dimensionally consistent.

Assume: $t =$ number $\times\ m^A l^B g^C$ sometimes written $t \propto m^A l^B g^C$

then: $[t] = [m^A l^B g^C]$ and $\mathsf{T} = \mathsf{M}^A \mathsf{L}^B (\mathsf{L}\mathsf{T}^{-2})^C = \mathsf{M}^A \mathsf{L}^B \mathsf{L}^C \mathsf{T}^{-2C} = \mathsf{M}^A \mathsf{L}^{B+C} \mathsf{T}^{-2C}$

so $\quad \mathsf{T} = \mathsf{M}^A \mathsf{L}^{B+C} \mathsf{T}^{-2C}$

For consistency we must have:

from M : $A = 0$

from T : $1 = -2C$

from L : $0 = B + C$

so $C = -\dfrac{1}{2}$ and $B = \dfrac{1}{2}$

Hence: $t =$ number $\times\ m^0 l^{\frac{1}{2}} g^{-\frac{1}{2}} =$ number $\times \sqrt{\dfrac{l}{g}}$ or $t \propto \sqrt{\dfrac{l}{g}}$

This is **not** a full proof of the pendulum formula and the method cannot be used to calculate the 2π constant in the pendulum formula. We shall give a more detailed discussion of this standard example in a later section.

Task

Use the method of dimensions to predict how:

(a) The tension T in a string may depend on the mass m of a particle which is being whirled round on its end in a circle of radius r and with speed v.

(b) The maximum height h reached by a stone may depend on its mass m, the energy E with which it is projected vertically and the acceleration due to gravity.

(c) The speed v of sound in a gas may depend on its pressure p, its density r and the acceleration g due to gravity.

Your solution

Answer

(a) $T \propto mv^2/r$ (b) $h \propto E/mg$ (c) $v \propto \sqrt{(p/r)}$

3. Finding the dimensional form of quantities

If we use the three basic dimensions mass, length and time, which we denote by the symbols M, L and T, then we shall need to find the dimensional formulae for quantities such as energy, angular momentum, etc., as we deal with a range of problems in mechanics. To do this we need to have a sufficient knowledge of the subject to enable us to recall important equations which relate the physical quantities which we are using in our mathematical theory. The dimensions of force and kinetic energy were derived briefly at the start of this section but to clarify the logical reasoning involved we now repeat the derivations in a slower step by step manner.

Perhaps the best-known equation in classical mechanics is the one expressing Newton's second law of particle motion (Force = mass × acceleration):

$$F = ma$$

To find the appropriate dimensional formula for force we can go through a sequence of operations as follows.

Step 1 Speed (or velocity) is usually quoted as so many metres per second or miles per hour and so its dimensional form is L/T or LT^{-1}.

Step 2 Acceleration is rate of change of velocity and so acceleration should be assigned the dimensional formula velocity/time, which leads to the result LT^{-2}.

Step 3 Using Newton's equation as given above we thus conclude that we obtain the expression for the dimensions of force by taking the product appropriate to mass × acceleration. Thus,

Force has the dimensions MLT^{-2}.

Example 1

Write down the formula for the kinetic energy of a particle of mass m moving with speed v and use it to decide what the dimensional expression for the energy should be.

Solution

The required formula is $E = \frac{1}{2}mv^2$ and so we obtain the dimensional expression for energy as:

Energy has the dimensions $\mathsf{M} \times (\mathsf{LT}^{-1})^2 = \mathsf{ML}^2\mathsf{T}^{-2}$.

4. Constants which have a dimensional expression

Several of the important equations which express physical laws include constants which have not only a numerical value (which can be looked up in a table of universal constants) but also have an appropriate dimensional formula. As a simple example, the "little g", g, used in the theory of falling bodies or projectiles might be given the value 32 feet per second per second or 981 cm per second per second (depending on the units used in our calculation). The shorthand terminology "32 feet per second squared" is often used in ordinary language and actually makes it clear that g is an acceleration, with the appropriate dimensional expression LT^{-2}.

Here are two examples of important equations which introduce constants which have both a numerical value and an associated dimensional expression. The first example is from quantum mechanics and the second is from classical mechanics.

1. In the quantum theory of light (relevant to the action of a modern laser) the energy of a light quantum for light of frequency ν is equal to $h\nu$, where h is called Planck's constant. This equation was formulated by Einstein in order to explain the behaviour of photoelectric emission, the physical process used in modern photoelectric cells. Planck's constant, h, which appears in this equation, later turned out to arise in much of quantum mechanics, including the theory of atomic structure.

2. The gravitational constant G of Newton's theory of universal gravitation appears in the equation which gives the gravitational force between two point masses M and m which are separated by a distance r:

$$F = \frac{GMm}{r^2}$$

Example 2

When quantum theory was applied to the theory of the hydrogen atom it was found that the motion of the electron around the nucleus was such that the angular momentum was restricted to be an integer multiple of the quantity $h/2\pi$, where h was the same Planck's constant which had arisen in the theory of radiation and of the photoelectric effect. Can you explain why this result is plausible by studying the dimensional formula for h?

Solution

We had earlier deduced that energy should be given the dimensional formula $\mathsf{ML}^2\mathsf{T}^{-2}$. Frequency is usually quoted as a number of cycles per second and so we should assign it the dimensional formula T^{-1} (on taking a pure number to be dimensionless and so to have a zero index for M, L and T). On taking h to be E/ν we conclude that h must be assigned the dimensional expression $\mathsf{ML}^2\mathsf{T}^{-2}/\mathsf{T}^{-1} = \mathsf{ML}^2\mathsf{T}^{-1}$. In particle mechanics the angular momentum is of the form "momentum $\times$ radius of orbit" i.e. mvr. This has the dimensions $\mathsf{M}(\mathsf{LT}^{-1})\mathsf{L} = \mathsf{ML}^2\mathsf{T}^{-1}$. Thus we see that Planck's constant, h, has the dimensions of an angular momentum. Dividing h by the purely numerical value 2π leaves this property unchanged. Indeed, if we rewrite the original equation $E = h\nu$ so as to use the angular frequency then we obtain $E = (h/2\pi)\omega$. The combination $h/2\pi$ is usually called "h bar" in quantum mechanics.

5. Further applications of dimensional analysis

Application 3: Deriving a formula

Almost every textbook dealing with dimensional analysis starts with the example of the simple pendulum and we did this in Application 2 a few pages back. We now look at this standard example in a slightly more general manner, using it to illustrate several useful methods and also to indicate the possible limitations in the use of dimensional analysis.

If we suppose that the periodic time t of a simple pendulum could possibly depend on the length l of the pendulum, the acceleration due to gravity g, and the mass m of the pendulum bob, then the usual approach is to postulate an equation of the form

$$t = \text{constant} \times m^A l^B g^C$$

Thus we concede from the start that the equation might include a multiplying factor (dimensionless) which is invisible to our approach via dimensional analysis. We now set out the related equation with the appropriate dimensional expression for each quantity

$$\mathsf{T} = \mathsf{M}^A \mathsf{L}^B (\mathsf{L}\mathsf{T}^{-2})^C = \mathsf{M}^A \mathsf{L}^{B+C} \mathsf{T}^{-2C}$$

We next look at M, L and T in turn and make their indices match up on the left and right of the equation. To do this we can think of the left-hand side as being $\mathsf{M}^0\mathsf{L}^0\mathsf{T}^1$. We find

(M) $\quad 0 = A$

(L) $\quad 0 = B + C$

(T) $\quad 1 = -2C$

giving $\quad A = 0,\ C = -1/2,\ B = 1/2.$

This leads to the conclusion that the periodic time takes the form $t = \text{constant} \times \sqrt{l/g}$.

This agrees with the result of simple mechanics (which also tells us that the numerical constant actually equals 2π). There are several points of interest here.

Comment 1

Although we were apparently looking at the problem of the simple pendulum our result simply shows that a time has the same dimensions as the square root of l/g. If we had started with the problem "If a particle starts from rest, how does the time t of fall depend on the distance fallen l, the gravitational acceleration g and the particle mass m?" we would have obtained the **same** result! The constant acceleration equations of simple mechanics give $l = \frac{1}{2}gt^2$ and so $t = \text{constant} \times \sqrt{l/g}$, where in this problem the constant has the value $\sqrt{2}$ instead of 2π.

Comment 2

The mass m does not enter into the final result, even though we included it in the list of variables which might possibly be relevant. With a more detailed knowledge of the problem we could have anticipated this by noticing that m does not appear in the differential equation which governs the motion of the pendulum:

$$l\frac{d^2\theta}{dt^2} = -g\sin(\theta)$$

On noting that the usual textbook theory replaces $\sin(\theta)$ by θ we arrive at Comment 3.

Comment 3

Our equation for t contains an unknown constant factor. This need **not** always be just a pure number, in the mathematical sense. For our pendulum problem we recall that the standard derivation of the equation for t involves the simplifying assumption that for a small angle θ we can set $\sin(\theta) = \theta$, where θ is the angle in radians. Thus we are admitting that the periodic time given by the usual textbook equation refers only to a **small** amplitude of swing. The basic equation which defines a radian is based on a ratio of distances; an angle of θ radians at the centre of a circle of radius r marks off an arc of length $r\theta$ along the circumference. Thus an angle in radians (or any other angular units) should be regarded as having dimensions $\mathsf{L}/\mathsf{L} = \mathsf{L}^0$ and is thus invisible in our process of dimensional analysis (just like a pure mathematical number). Thus for our simple pendulum problem we could set

$$t = F(\theta_0)\sqrt{l/g}$$

where F is some function of the amplitude θ_0 of the swing. Our dimensional analysis would not be able to tell us about F, which is invisible to the analysis. If we think about the physics of the problem then we can see that t should decrease with θ_0 and should be an even function (since θ_0 and $-\theta_0$ represent the same motion). We can thus reasonably suppose that the form of $F(\theta_0)$ starts off as $2\pi(1 - A\theta_0^2 \ldots)$, with A being a positive number. However, this detail is not visible to our dimensional analysis. This example highlights the fact that dimensional analysis is only an auxiliary tool and must always be backed up by an understanding of the system being studied. This need for a preliminary understanding of the system is evident right at the start of the analysis, since it is needed in order to decide which variables should be included in the list which appears in the postulated equation. An approach to the simple pendulum problem using a conservation of energy approach can be used; it leads to an expression for the periodic time which does indeed show that the periodic time is an even function of the amplitude θ_0 but it involves the evaluation of elliptic integrals, which are outside the range of this workbook.

Comment 4

Another way of writing the first simple result for the periodic time of the pendulum would be to use a dimensionless variable:

$$gt^2/l = \text{constant}$$

We now know that strictly speaking the right-hand side of this equation need be a constant only to the extent that it is dimensionless (so that it could depend on a dimensionless quantity such as the amplitude of the swing). The approach of using dimensionless variables to represent the equations governing the behaviour of a system is often useful in modelling experiments and is particularly associated with the work of Buckingham. The Buckingham Pi theorem is quoted in various forms in the literature. Its essential content is that if we start with n variables then we can always use our basic set of dimensional units (M, L and T in the pendulum case) to construct a set of dimensionless variables (fewer than n) which will describe the behaviour of the system. This idea is used in some of our later examples.

Planetary motion: Kepler's laws

Newton showed that the combination of his law of universal gravitation with his second law of motion would lead to the prediction that the planets should move in elliptic orbits around the Sun. Further, since the gravitational force is central (i.e. depends only on r and not on any angles) it follows that a planet moves in a plane and has constant angular momentum; this second fact is equivalent to the Kepler law which says that a planet "sweeps out equal areas in equal times". Another Kepler law relates the time t for a planet to go around the Sun to the mean distance r of the planet from the Sun (or rather to the major semi-axis of its elliptic orbit). To attack this problem using dimensional analysis we need to find the dimensional formula to assign to the gravitational constant G. We start from the equation given on page 9 for the inverse square law force and rewrite it to give

$$G = \frac{Fr^2}{mM}$$

So that we arrive at the dimensional formula associated with G:

$$G \textbf{ has the dimensions } (\mathsf{MLT}^{-2})\mathsf{L}^2\mathsf{M}^{-2} = \mathsf{ML}^3\mathsf{T}^{-2}.$$

Example 3

Find how the periodic time t for a planetary orbit is related to the radius r of the orbit.

Solution

We suppose that t depends on G, on the Sun's mass M and on r. (Looking at the planet's equation of motion suggests that the planet's mass m can be omitted). We set

$$t = \text{constant} \times G^A M^B r^C$$

leading to the dimensional equation

$$\mathsf{T} = (\mathsf{M}^{-1}\mathsf{L}^2\mathsf{T}^{-2})^A \mathsf{M}^B \mathsf{L}^C$$

Comparing coefficients on both sides gives three equations:

(M) $\quad 0 = B - A$

(L) $\quad 0 = 3A + C$

(T) $\quad 1 = -2A$

which produces the result $A = -1/2,\ B = -1/2,\ C = 3/2$.

If we use t^2 instead of t we conclude that t^2 is proportional to $\dfrac{r^3}{GM}$. This is in accord with the Kepler law which states that t^2/r^3 has the same value for all the planets.

The search title "Kepler's laws" in Google gives sites with much detail about the motion both of planets around the Sun and of satellites around the planets. The ratio t^2/r^3 varies by about one percent across all the planets. (See the following Comment)

Comment

You might have noticed that we used two laws due to Newton in the calculation above but ignored his third law which deals with action and reaction. The Sun is acted on by an equal and opposite force to that which it exerts on the planet. Thus the Sun should move as well! In principle it goes round a very small orbit. A detailed mathematical analysis shows that for the special case of one sun of mass M and one planet of mass m then the correct mass to use in the calculation above is the reduced mass given by the formula $mM/(m+M)$. This has the dimensions of a mass and thus our simple dimensional analysis cannot distinguish it from M. The modified Kepler's law then predicts that r^3/t^2 should be proportional to $(M+m)$ and the planets obey this modified law almost perfectly.

This again illustrates the need for as much detailed knowledge as possible of a system to back up the use of dimensional analysis. In the real solar system, of course, there are several planets and they exert weak gravitational forces on one another. Indeed the existence of the planet Neptune was surmised because of the small deviations which it produced in the motion of the planet Uranus.

An example from electrostatics and atomic theory

In the electrostatic system of units, the force F acting between two charges q and Q which are a distance r apart is given by an inverse square law of the form

$$F = \text{constant} \times \frac{qQ}{r^2}, \qquad \text{where the constant} = 1 \text{ or } 4\pi \text{ depending on the units used.}$$

If we wish to decide on an appropriate dimensional formula to represent charge then we set $Q = \mathsf{M}^A\mathsf{L}^B\mathsf{T}^C$ in the dimensional equation which represents this inverse square law equation. We have

$$\mathsf{MLT}^{-2} = (\mathsf{M}^A\mathsf{L}^B\mathsf{T}^C)^2\mathsf{L}^{-2}$$

Comparing coefficients on both sides gives the equations

(M) $\quad 1 = 2A$

(L) $\quad 1 = 2B - 2$

(T) $\quad -2 = 2C$

Thus $A = 1/2,\ B = 3/2,\ C = -1$, so Q^2 has the dimensional formula $\mathsf{ML}^3\mathsf{T}^{-2}$. ($Q^2$ is simpler to use than Q and appears in most other results.)

Example 4

In the theory of the hydrogen atom the energy levels of the atom depend on the principal quantum number $n = 1, 2, 3, \ldots$ and equal $\dfrac{-1}{2n^2}$ times an energy unit which depends on the electronic charge e, the electron mass m and Planck's constant h (which we have seen in an earlier problem to have the same dimensions as an angular momentum). Use dimensional analysis to find how the atomic energy unit (the Rydberg) depends on $e,\ m$ and h.

Solution

We require that the combination $e^{2A}m^Bh^C$ should have the dimensions of an energy.

In terms of dimensional formulae we thus obtain the equation

$$\mathsf{ML^2T^{-2}} = (\mathsf{ML^3T^{-2}})^A\mathsf{M}^B(\mathsf{ML^2T^{-1}})^C$$

Comparing coefficients on both sides gives the equations

(M) $\quad 1 = A + B + C$

(L) $\quad 2 = 3A + 2C$

(T) $\quad -2 = -2A - C$

Adding twice the (T) equation to the (L) equations shows that $A = 2$ and then the other variables are found to be $B = 1$ and $C = -2$. Thus we find that the Rydberg atomic energy unit must depend on $e,\ m$ and h via the combination me^4/h^2, which is as given by the full theory. Indeed, the quantities $e,\ m$ and h are the quantities appearing in the Schrödinger equation for the hydrogen atom and are thus the appropriate ones to use.

Some examples from the theory of fluids

In fluid theory the concepts of viscosity and of surface tension play an important role and so we must decide on the appropriate dimensional formula to assign to the quantities which are associated with these concepts. The coefficient of viscosity μ of a fluid is such that the force per unit area across a fluid plane is μ times the fluid velocity gradient perpendicular to that surface. (We note that it is an assumption based on a combination of theory and experience to assert that such a well defined coefficient does exist; in general a fluid is called a Newtonian fluid if it has a well defined μ. Even so, μ can vary with temperature and other variables).

If we set force/area $= \mu \times$ velocity/distance then we find

$$\mu = (\text{force} \times \text{distance})/(\text{area} \times \text{velocity}) = (\mathsf{MLT^{-2}} \times \mathsf{L})/(\mathsf{L^2} \times \mathsf{LT^{-1}}) = \mathsf{ML^{-1}T^{-1}}$$

Example 5

Suppose that a sphere of radius r moves at a slow speed v through a fluid with coefficient of viscosity μ and density ρ. Find the way in which the resistive force acting on the sphere should depend on these three quantities in two different ways:

(a) by omitting ρ from the list of variables

(b) by omitting μ from the list of variables

Solution

(a) We set $F = \text{constant} \times \mu^A r^B v^C$ and so arrive at the dimensional equation

$$\mathsf{MLT}^{-2} = (\mathsf{ML}^{-1}\mathsf{T}^{-1})^A \mathsf{L}^B (\mathsf{LT}^{-1})^C$$

Comparing coefficients on both sides gives

(M) $1 = A$

(L) $1 = B + C - A$

(T) $-2 = -A - C$

which yields the result $A = 1,\ B = 1,\ C = 1.$

Thus the resistance is proportional to the product $rv\mu$. Detailed theory gives this result (called Stokes' law) with the constant equal to 6π. This formula is appropriate for very slow motion.

We now include the density ρ in the analysis and omit μ.

(b) On supposing that the resistive force is given by $F = r^A v^B \rho^C$ we obtain the dimensional equation

$$\mathsf{MLT}^{-1} = (\mathsf{L})^A (\mathsf{LT}^{-1})^B (\mathsf{ML}^{-3})^C$$

leading to the three equations:

(M) $1 = C$

(L) $1 = A + B - 3C$

(T) $-2 = -B$

These equations give the results $A = 2, B = 2, C = 1$, so that the force is calculated to vary as $\rho r^2 v^2$. This expression for the resistance is appropriate for speeds which are not too high but at which the moving object has to push the fluid aside (thus giving it some energy related to ρv^2) while the Stokes' force of Solution (a) refers to motion so slow that this dynamical effect is negligible. In the theory of flight the force on an aircraft's wing has two components, lift and drag. These depend on the aircraft's speed and on the angle of inclination of the wing. The expression $\rho r^2 v^2$ can be interpreted as being area$\times\rho v^2$ for a wing with a given surface area and the numerical multiplying factor which gives the actual drag and lift as a function of the wing angle of inclination is expressed as a lift or drag coefficient.

Comment

In Solutions (a) and (b) we have obtained two different expressions with the dimensions of a force, by selecting three variables at a time from a list of four variables. We can clearly obtain a dimensionless quantity by taking the ratio of the two different force expressions. This gives us

$$\rho r^2 v^2/(r\mu v) = \rho r v/\mu$$

This dimensionless quantity is called the Reynold's number and it appears in various parts of the theory of fluids. In this example r was the radius of a moving spherical object; in other cases it can be the radius of a pipe through which the fluid is flowing. Experiments show that (in both cases) the fluid flow is smooth at low Reynold's numbers but becomes turbulent above a critical Reynold's number (the numerical value of which depends on the particular system being studied). If we adopt the point of view which writes the results of dimensional analysis in the form of equations which involve dimensionless variables then the results of problems (a) and (b) can be written in a form which involves the Reynold's number:

$$F/(\rho v^2 r^2) = f(\rho r v/\mu)$$

Where f is some function of the Reynold's number. If $f(x)$ starts off as $6\pi/x$ and tends to a constant for large v then we will obtain the two special cases (a) and (b) which we treated in the problems. If you are familiar with the history of flight you might know that if v is sufficiently large for the compressibility of the fluid to play an important role then f should be extended to involve another dimensionless variable, the Mach number, equal to v/v_s where v_s is the speed of sound.

Example 6

Suppose that a fluid has the viscosity coefficient μ and flows through a tube with a circular cross section of radius r and length l. If the pressure difference across the ends of the tube is p, find a possible expression for the volume per second of fluid passing through a tube, using pressure gradient p/l as a variable.

Solution

We set volume per second = constant $\times\, r^A \times \left(\dfrac{p}{l}\right)^B \times \mu^C$ and obtain the dimensional equation

$$\mathsf{L}^3\mathsf{T}^{-1} = \mathsf{L}^A(\mathsf{M}\mathsf{L}^{-2}\mathsf{T}^{-2})^B(\mathsf{M}\mathsf{L}^{-1}\mathsf{T}^{-1})^C$$

leading to the equations

(M) $\quad 0 = B + C$

(L) $\quad 3 = A - 2B - C$

(T) $\quad -1 = -2B - C$

The result is found to be $A = 4, B = 1, C = -1$, thus leading to the conclusion that the volume per second should vary as $r^4\mu^{-1}pl^{-1}$. This is the result given by a theoretical treatment which allows for the fact that the fluid sticks to the walls and so has a speed which increases from zero to a maximum as one moves from the wall to the centre of the tube.

A problem involving surface tension

If we recall that the surface tension coefficient σ can be regarded as an energy per unit area of a liquid surface then we can conclude that σ has the dimensions $\mathsf{ML}^2\mathsf{T}^{-2}/\mathsf{L}^2 = \mathsf{MT}^{-2}$.

Example 7

Consider a wave of wavelength λ travelling on the surface of a fluid which has the density ρ and the surface tension coefficient σ. Including the acceleration due to gravity g in the list of relevant variables, derive possible formulae for the velocity of the wave as a function of λ by

(a) omitting σ from the list, (b) omitting g from the list.

Solution

(a) The equation $v =$constant $\times\ \lambda^A\rho^B g^C$ leads to the dimensional equation

$$\mathsf{LT}^{-1} = \mathsf{L}^A(\mathsf{ML}^{-3})^B(\mathsf{LT}^{-2})^C$$

which gives the linear equations

(M) $0 = B$

(L) $1 = A - 3B + C$

(T) $-1 = -2C$

The solution is quickly seen to be $A = \frac{1}{2}, B = 0, C = \frac{1}{2}$. We find that v^2 is proportional to $g\lambda$.

(b) If we set $v =$constant $\times\ \lambda^A\sigma^B\rho^C$ then we obtain the dimensional equation

$$\mathsf{LT}^{-1} = \mathsf{L}^A(\mathsf{MT}^{-2})^B(\mathsf{ML}^{-3})^C$$

leading to the linear equations

(M) $0 = B + C$

(L) $1 = A - 3C$

(T) $-1 = -2B$

This gives the solution $A = -\frac{1}{2}, B = \frac{1}{2}, C = -\frac{1}{2}$. We find that v^2 should be proportional to $\sigma/(\lambda\rho)$.

We note that the results (a) and (b) depend on the wavelength λ in different ways, which might suggest that (a) is more relevant for long waves and (b) is more relevant for short waves. Detailed solution of the fluid equations of motion gives the result

$$v^2 = g\lambda/(2\pi) + 2\pi\sigma/(\lambda\rho)$$

showing that long waves involve a kind of up and down oscillation of the water mass under the action of gravity while short waves involve a strong curvature of the surface which brings surface tension forces into play.

Comment

The preceding examples dealing with fluids both involve situations which are more complicated than the simple pendulum problem which was treated earlier. In particular, both the fluid problems are such that if we use the three basic dimensions M, L and T then we can form **two** different dimensionless quantities from the list of physical variables which are regarded as important for the problem being studied. Thus, if we take the traditional simple approach we can obtain two different possible results by removing one variable at a time. If we use the Buckingham approach based on dimensionless variables then we arrive at an equation which says that the first dimensionless quantity is some arbitrary function of the second one.

An example involving temperature

The van der Waals' equation of state was an early equation which tried to describe the way in which the behaviour of a real gas differs from that of an ideal gas. One mole of an ideal gas obeys the equation $PV = RT$, where P, V and T are the pressure and volume of the gas and T is the absolute temperature. R is the universal gas constant. The simple equation of state can be derived from kinetic theory by assuming that the gas atoms are of negligible size and that they move independently. van der Waals' equation introduced two parameters a and b into the equation of state; his modification of the ideal gas equation takes the form

$$(P + aV^{-2})(V - b) = RT$$

The parameter a is intended to allow roughly for the effects of interactions between the atoms while b allows for the finite volume of the atoms. When we attempt to apply the methods of dimensional analysis to this equation we run into a problem of notation. We have already used the symbol T to represent time and so must not confuse that symbol with the temperature T in the traditional theory of gases. Since the absolute temperature is often quoted as "degrees Kelvin" we shall adopt the symbol K for temperature. The product PV in the ideal gas equation has the dimensions of an energy, since (force per unit area) x (volume) = force x distance = energy. If we now use a symbol K for a new dimensional quantity, the temperature, then we conclude that the gas constant R is a dimensional constant, with the dimensional formula given by

$$[\text{Energy}/K] = \mathsf{ML^2T^{-2}K^{-1}}$$

Inspection of the van der Waals' equation makes it clear that if we re-arrange it so as to find the volume V of the gas at a given pressure and temperature then we arrive at a **cubic** equation for V. For some P and T values this equation has one real and two complex conjugate roots and so it is reasonable to take the real root as being the one which represents the actual physical volume of the gas. However, there are regions in which the cubic equation has three real roots, which sets us a problem of interpretation! Various workers have tackled this problem by trying to apply the principles of thermodynamics (in particular, in attempting to find the solution with the minimum free energy) and their conclusion is that the "triple real solution" region is actually one in which the gas is condensing into a liquid (so that the volume contains both gas and liquid portions). If the temperature is sufficiently high, above what is called the critical temperature, then the gas cannot be made to condense even when the pressure is increased.

Example 8

Work out the dimensional formulae for the parameters a and b in the van der Waals' equation and then find a possible equation which gives the critical temperature of the gas in terms of a, b and R.

Solution

In van der Waals' equation $(P + aV^{-2})(V - b) = RT$ the term aV^{-2} is in the bracket with the pressure P and so must itself have the dimensions of a pressure i.e. a force per unit area. Thus we conclude that a has the dimensions of
[force] x [squared volume]/[area], which gives the result

$$(\mathsf{MLT}^{-2}) \times \mathsf{L}^6/\mathsf{L}^2 = \mathsf{ML}^5\mathsf{T}^{-2}$$

b clearly has the dimensions L^3 of a volume. We have already obtained the dimensions of R in our previous discussion and so we try the following expression for the critical temperature:

$$K = a^A b^B R^C$$

The dimensional equation corresponding to our proposed equation is

$$\mathsf{K} = (\mathsf{ML}^5\mathsf{T}^{-2})^A(\mathsf{L}^3)^B(\mathsf{ML}^2\mathsf{T}^{-2}K^{-1})^C$$

This gives four linear equations

(M) $\quad 0 = A + C$

(L) $\quad 0 = 5A + 3B + 2C$

(T) $\quad 0 = -2A - 2C$ (this is the same as the M equation)

(K) $\quad 1 = -C$

There thus turn out to be only three independent equations and these have the solution

$$A = 1, B = -1, C = -1.$$

Thus we conclude that the critical temperature should be a constant times $a/(bR)$, as is concluded by the detailed theory, which gives the value $8/27$ for the constant. The critical temperature turns out to be the temperature associated with the special P-V trajectory along which the cubic equation for V has a triple real root at one point, which is called the triple point.

6. Water flowing in a pipe

Introduction

One way of measuring the flow rate of water in a pipe is to measure the pressure drop across a circular plate, containing an orifice (hole) of known dimensions at its centre, when the plate is placed across the pipe. The resulting device is known as an orifice plate flow meter. The theory of fluid mechanics predicts that the volumetric flow rate, Q (m^3 s^{-1}), is given by

$$Q = C_d A_o \sqrt{\frac{2g\Delta h}{1 - A_o^2/A_p^2}}$$

where Δh (m) is the pressure drop (i.e. the difference in 'head' of water across the orifice plate) A_o (m^2) and A_p (m^2) are the areas of the orifice and the pipe respectively, g is the acceleration due to gravity (m s^{-2}) and C_d is a discharge coefficient that depends on viscosity and the flow conditions.

Problem in words

(a) Show that C_d is dimensionless.

(b) Rearrange the equation to solve for the area of the orifice, A_o, in terms of the other variables.

(c) Calculate the orifice diameter required if the head difference across the orifice plate is to be 200 mm when a volumetric flow rate of 100 cm^3 s^{-1} passes through a pipe with 10 cm inside diameter. Assume a discharge coefficient of 0.6 and take $g = 9.81$ m s^{-2}.

Mathematical analysis

(a) The procedure for checking dimensions is to replace every quantity by its dimensions (expressed in terms of length L, time T and mass M) and then equate the powers of the dimensions on either side of the equation. Q, the volume flow rate has dimensions m^3 s^{-1} which can be written L^{3}T^{-1}. Areas A_o and A_p have dimensions of m^2 which can be written L^2. g has dimensions m s^{-2} which can be written L T^{-2}. None of these involve the mass dimension. Consider the square root on the right-hand side of the equation. First look at the denominator. Since areas A_o and A_p have the same dimensions, A_o^2/A_p^2 is dimensionless. The number 1 is dimensionless, so the denominator is dimensionless. Now consider the numerator. The product $g\Delta h$ has dimensions LT^{-2}L = L^{2}T^{-2} . This means that the square root has dimensions LT^{-1}. After multiplying by an area (A_o) the dimensions are L^{3}T^{-1}. These are the dimensions of Q. So the factors other than C_d on the right-hand side of the equation together have the same dimensions as Q. This means that C_d must be dimensionless.

(b) Starting from the given equation $\quad Q = C_d A_o \sqrt{\dfrac{2g\Delta h}{1 - A_o^2/A_p^2}}$

Square both sides: $\quad Q^2 = C_d^2 A_o^2 \dfrac{2g\Delta h}{1 - A_o^2/A_p^2}$

Now multiply through by $1 - A_o^2/A_p^2$. $\quad Q^2(1 - A_o^2/A_p^2) = 2g\Delta h C_d^2 A_o^2$

Collect together terms in A_o^2: $\quad Q^2 = A_o^2(2g\Delta h C_d^2 + Q^2/A_p^2)$

Divide through by the bracketed term on the right-hand side of this equation and square root both sides:

$$A_o = \frac{Q}{\sqrt{2g\Delta h C_d^2 + Q^2/A_p^2}}$$

(c) Substitute $\Delta h = 0.2,\ A_p = 0.1,\ g = 9.81,\ Q = 10^{-3}$ and $C_d = 0.6$ in the above to give $A_o = 1.189 \times 10^{-3}$. The required orifice area is about 1.2×10^{-3} m^2.

7. Temperature as a dimensional variable

In Example 8 we introduced temperature as a new independent dimensional variable. The universal gas constant R will then acquire the dimensions of $\dfrac{[PV]}{[K]} = \dfrac{[\text{Energy}]}{[K]} = \text{M L}^2\ \text{T}^{-2}\ \text{K}^{-1}$, and other quantities in thermodynamic theory will acquire corresponding dimensional formulae.

Temperature is one of the fundamental entities on the International Units System, along with mass, length and time, which have played the major role in the examples which we have discussed. The standard symbol θ is often used for temperature, although we used K, since θ had been used to denote an angle in several examples. Example 4 involved a combination of electrostatics and dynamics and for the particular calculation treated it was sufficient to describe the electrical charge Q in terms of M, L and T. However, the International Units System has electrical current I as a basic dimensional quantity, so that charge Q becomes current times time, i.e. with dimensions IT. The potential difference between two points, being energy (i.e. work done) per unit charge would have dimensions given by $\text{ML}^2\ \text{T}^{-2}\ (\text{IT})^{-1} = \text{ML}^2\ \text{I}^{-1}\ \text{T}^{-3}$, whereas in the approach used in our Example 4 it would have the dimensions $\text{M}^{\frac{1}{2}}\ \text{L}^{\frac{1}{2}}\ \text{T}^{-1}$, since current would have had the dimensions $\text{M}^{\frac{1}{2}}\ \text{L}^{\frac{3}{2}}\ \text{T}^{-2}$. It is clear that in the (MLTI) system the resulting expressions are simpler than they would be if we attempted to do every calculation by reducing everything to the (MLT) system.

One basic difference between the approach of Example 4 and that which is adopted in Electrical Engineering (with I as a fundamental dimension) arises from the equation for the law of force between two charges. We took the force to be equal to a constant times the product of the two charges divided by the distance squared, whereas in modern approaches to electromagnetic theory our "constant" is taken to be a dimensional quantity (the reciprocal of the permittivity of free space).

Historically, the need for a set of dimensional quantities and a set of units which were jointly appropriate for both electrical and magnetic phenomena became more obvious with the publication of Maxwell's equations and the discovery of electromagnetic waves, in which the electrical and magnetic fields propagate together. This topic is too lengthy to be treated here but is discussed in detail in some of the works cited in subsection 8.

8. Some useful references

Nowadays there are many articles, books and web pages which deal with dimensional analysis and its applications.

8.1 Three Classic Books

(1) The book *Dimensional Analysis* by P.W. Bridgman is regarded as probably the most careful study of the power (and of the limitations) of dimensional analysis. It gives many applications but also stresses the importance of starting from a detailed knowledge of the governing equation of motion (often a differential equation) for a system in order to be able to make an intelligent choice of the variables which should appear in any expression which is to be analysed using dimensional analysis.

(2) The book *Hydrodynamics* by G. Birkhoff adopts a similar critical approach but deals specifically with problems involving fluid flow (where dimensionless quantities such as the Reynold's number play a role). Birkhoff uses some concepts from group theory in his discussion and gives the name "inspectional analysis" to the total investigative approach which combines dimensional analysis with other mathematical techniques and with physical insight in order to study a physical system.

(3) The book *Method of Dimensions* by A.W. Porter contains many examples from mechanics, thermodynamics and fluid theory, with discussion of both the power and the limitations of dimensional analysis.

Book (1) was published by Yale University press, (2) by Oxford University Press (in Britain) and (3) by Methuen. There should be a copy in most university libraries.

8.2 Some journal articles

The American journal of Physics has published several detailed explanations of the use of dimensional analysis together with particular examples. Here are a few:

- Volume 51, (1983) pages 137 to 140 by W.J. Remillard. *Applying Dimensional Analysis.*
- Volume 53, (1985) pages 549 to 552 by J.M. Supplee. *Systems of equations versus extended reference sets in dimensional analysis.* (This mentions the possibility of using separate L_x and L_y dimensions to replace the single L dimension for some special cases when the problem involves motion in two dimensions.)
- Volume 72, (2004) pages 534 to 537 by C.F. Bohren. *Dimensional analysis, falling bodies and the fine art of not solving differential equations.* (This gives a lengthy study of the theory of a body falling in the Earth's gravitational field, gradually introducing the effects of the Earth's finite radius, the Earth's rotation and the presence of air resistance.)
- Volume 71, (2003) pages 437 to 447 by J.F. Pryce. *Dimensional analysis of models and data sets.* (This article deals with the simple pendulum, with air resistance being discussed and with the example being used to illustrate general ideas about modelling and about the use of dimensionless variables.)

8.3 Web Sites

The site www.pigroup.de connects with a research group at the University of Stuttgart which uses dimensional analysis in various engineering applications. The site gives a list of the published work of the group.

8.4 Using Google

The simplest way to search nowadays is to type in an appropriate title in the Google search engine. To explore particular topics treated or mentioned in this section it is best to use titles such as:

Buckingham Pi Theorem (to see the different interpretations of this theorem).

Reynold's Number (to see the origin and uses of this number).

Wind Tunnel (to see details of modelling projects using dimensional analysis in the study of air resistance for various objects, including tennis balls).

Drag Coefficient (to see how the air resistance to a wing varies as the speed increases with particular reference to the lift and drag forces on a wing).

Van der Waals' Equation + **Critical Point** (this gives several pages which explain the theory of the van der Waals' equation and discuss how well it describes the behaviour of real gases; several pages also give alternative equations proposed by other workers).

Wikipedia Potential Difference

The topic touched on in Section 7, the use of the current I as a fundamental quantity in the dimensional analysis of problems involving electricity and magnetism, is an important one and so appears on several websites. The Google input Wikipedia Potential Difference leads to a useful table which sets out the dimensional expressions for many common quantities in terms of the (MLTI) system. The history of the development of the system of fundamental dimensions in electromagnetism is explained in the article by G. Thomas (Physics Education, Volume 14, page 116, 1979); this article is also obtained via Google.

You can, of course, try any other search title, if you have the patience to keep going until you find a particular site which gives you something useful!

Mathematical Explorations

47.2

Introduction

This Section revisits some of the mathematics already introduced in other Workbooks but explores several alternative ways of looking at it, with the intention of widening the range of techniques or approaches which the student can call on when dealing with mathematical problems. Familiar functions such as the tangent function of trigonometry are used to illustrate the methods which are introduced, with an application to optics being used to motivate the study of the mathematics involved. Several links between apparently different pieces of mathematics are introduced to encourage the student to appreciate the value of looking for underlying structures which will permit 'knowledge transfer' between what at first sight appear to be topics from different Workbooks.

Prerequisites

Before starting this Section you should . . .

- be familiar with the concept of the Maclaurin series and the method of comparing coefficients of powers of x for power series
- know and understand the definitions of the trigonometric functions
- know and be able to use the theorem of Pythagoras
- know the basic theory of geometric series

Learning Outcomes

On completion you should be able to . . .

- appreciate the value of moving between different approaches to a problem e.g. from a trigonometric diagram to algebra or from a complex number calculation to a matrix calculation
- use a power series approach to problems in order to supplement or replace alternative calculations based on traditional algebra
- have some appreciation of the idea of an isomorphism in mathematics and how this can be useful in extending the range of applicability of a mathematical technique

Introduction to the exploratory mathematical examples

The $\tan(x)$ function is usually regarded as the most "difficult" of the three standard trigonometric functions. Examples 1 to 4 show that it has interesting properties and that it has uses in coordinate geometry and in optics. Example 5 deals with approximating a parabola by an arc of a circle. Example 6 explores efficient ways to find Maclaurin series. Example 7 demonstrates a surprising link between complex arithmetic and matrix algebra. Examples 8 and 9 provide unusual approaches to the exponential function, and finally Example 10 derives the link between the inverse hyperbolic tangent and natural logarithms.

Example 1

Maclaurin series of $\tan x$

Derive the first four non-zero terms in the Maclaurin series of $\tan(x)$ by using the fact that $\tan(x)$ has the derivative $1 + \tan^2(x)$.

Solution

First we note that the series has only odd powers of x in it because $\tan(x)$ is an odd function, with the property $f(-x) = -f(x)$. Thus we have the Maclaurin series

$$\tan(x) = A_1x + A_3x^3 + A_5x^5 + \ldots$$

Using the known result $T' = 1 + T^2$ we must have

$$\begin{aligned} A_1 + 3A_3x^2 + 5A_5x^4 + 7A_7x^6 \ldots &= 1 + (A_1x + A_3x^3 + A_5x^5 \ldots)^2 \\ &= 1 + A_1^2x^2 + 2A_1A_3x^4 + (A_3^2 + 2A_1A_5)x^6 + \ldots \end{aligned}$$

Comparing the coefficients of x^N on both sides gives the results

$(N = 0)$ $A_1 = 1$

$(N = 2)$ $3A_3 = A_1^2$, so that $A_3 = 1/3$

$(N = 4)$ $5A_5 = 2A_1A_3 = 2/3$, so that $A_5 = 2/15$

$(N = 6)$ $7A_7 = A_3^2 + 2A_1A_5 = 1/9 + 4/15 = 17/45$, so that $A_7 = 17/315$

We have thus found the expansion up to the x^7 term.

A careful study of the structure of the equations used above shows us how to obtain a general equation which describes the process. The general recurrence relation for the calculation is

$$(N+1)A_{N+1} = \sum_{J=1}^{N-1} A_J A_{N-J}$$

This can easily be programmed for a computer to generate many terms of the series.

Example 2

Derivatives of $\tan x$

When higher derivatives are being used in the differential calculus the symbol $F^{(n)}$ is often used to denote the n^{th} derivative of a function F. For the function $T = \tan(x)$ we know that $T^{(1)} = 1 + T^2$. Use the chain rule to work out the derivatives of $\tan(x)$ up to $T^{(5)}$ and show how this can give the first three terms in the Maclaurin series for $\tan(x)$.

Solution

By the chain rule we have

$$T^{(n+1)} = \frac{d}{dx}(T^{(n)}) = \frac{d}{dT}(T^{(n)}) \times \frac{dT}{dx} = (1+T^2)\frac{d}{dT}(T^{(n)})$$

Thus given the result

$$T^{(1)} = 1 + T^2$$

we find that

$$T^{(2)} = (1+T^2)\frac{d}{dT}(T^{(1)}) = (1+T^2)2T = 2T + 2T^3$$

$$T^{(3)} = (1+T^2)(2+6T^2) = 2 + 8T^2 + 6T^4$$

$$T^{(4)} = (1+T^2)(16T + 24T^3) = 16T + 40T^3 + 24T^5$$

$$T^{(5)} = (1+T^2)(16 + 120T^2 + 120T^4) = 16 + 136T^2 + 240T^4 + 120T^6$$

To find the Maclaurin series by the textbook method we need to find the n^{th} derivative of $\tan(x)$ at $x = 0$ and then divide it by n factorial $(n!)$. $T = 0$ at $x = 0$ and so the required n^{th} derivative is just the constant term in the $T^{(n)}$ calculated above. This term is zero for the even n and leads to the correct first three terms of the $\tan(x)$ series when the terms with odd n are taken. If we set down the general equation relating derivatives of T to powers of T,

$$T^{(n)} = \sum A(n,m)T^m$$

and insert this in the chain rule equation used in the calculations above, then we arrive at a recurrence relation by comparing the T^m coefficients on both sides:

$$A(n+1,\ m) = (m+1)\ A(n, m+1) + (m-1)\ A(n, m-1)$$

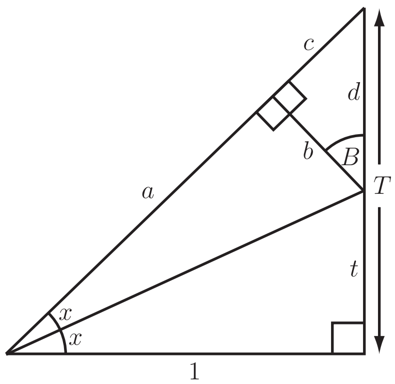
which can be used in a computer calculation or to set up a tabular method of calculating the results above, rather like a more complicated version of Pascal's triangle. Here are the first few rows:

$n \setminus m$	0	1	2	3	4	5
1	1	0	1	0	0	0
2	0	2	0	2	0	0
3	2	0	8	0	6	0
4	0	16	0	40	0	24

Example 3

A geometric derivation of the equation for tan($2x$)

To study the reflections from a parabolic mirror (Example 4) we need to know how to find $\tan(2x)$ from $\tan(x)$. This Example presents a simple way to derive the required formula, while giving some practice at algebraic manipulation and simple geometry. Study the following diagram, in which t denotes $\tan(x)$ and T denotes $\tan(2x)$. Find the lengths a, b, c and the angle B on the diagram and thence derive an equation for T by studying the small triangle with the sides $b,\ c$ and d. You will need to do some algebra to work the equation into a form where it gives T in terms of t. When you have obtained the equation, use it to include the length d in your calculation and so find an equation which gives $\cos(2x)$ in terms of t. (We are in fact using a geometrical approach to find what are called "the tan half angle formulae" in trigonometry.)



Solution

By studying the two similar triangles we see that $a = 1$ and $b = t$. Using the fact that the angles in a plane triangle add up to $180°$, we find that the angle B is just $2x$, which means that $T = \tan(2x)$ must equal c/b. To find c we have to note that $a + c$ is the hypotenuse of a triangle in which the other two sides have the lengths 1 and T. We can thus use the theorem of Pythagoras to find $a + c$. Putting all these facts together leads to an equation:

$$T = c/b = \left(\sqrt{1 + T^2} - 1\right)/t$$

This leads to $Tt + 1 = \sqrt{1 + T^2}$. Squaring gives $T^2t^2 + 2Tt + 1 = 1 + T^2$

Cancelling the 1 on both sides and dividing by the non-zero number T gives, after re-arranging terms

$$T = 2t/(1 - t^2), \quad \text{so that} \quad \tan 2x = \frac{2\tan x}{(1 - \tan^2 x)}.$$

If we now look at the length d we find that it equals $T - t$, so that $\cos(2x) = b/d = t/(T - t)$

Inserting the result for T in this equation and tidying up the fractions which appear leads to the results

$$\cos(2x) = (1 - t^2)/(1 + t^2), \quad \sin(2x) = 2t/(1 + t^2)$$

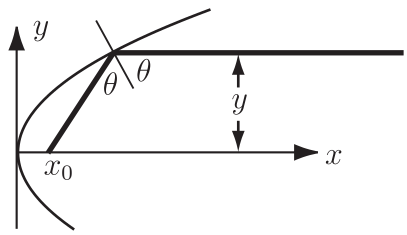
where the result for $\sin(2x)$ follows since $\sin(2x) = \tan(2x)\cos(2x)$.

Example 4

An application of the tan(2x) equation to optics

One of the useful applications of the tangent function arises in coordinate geometry. If a line has the equation $y = mx + c$ then the angle θ which the line makes with the x-axis satisfies $\tan(\theta) = m$. In dealing with optical reflections we know that the angle of incidence equals the angle of reflection. The angle between the incident and reflected light beams is thus exactly twice the angle of incidence. This problem exploits that fact to make use of the double angle equation derived in Example 3.

Consider a parabola which is described by the equation $y = A\sqrt{x}$. Suppose that a beam of light travelling horizontally from right to left hits the parabola at a height y above the x-axis and that it undergoes a mirror reflection from the parabola. Find the x coordinate at which the light beam crosses the x-axis and show that it is independent of the value of y. The diagram below shows the details.



Solution

The gradient at the point (x, y) on the parabola is $\dfrac{dy}{dx} = \dfrac{1}{2}Ax^{-1/2}$. This is more conveniently written as $(1/2)A^2/y$. The normal to the parabola has the gradient $-2y/A^2$, since two lines at right-angles have the property that the product of their gradients is -1. This result means that the normal makes an angle θ with the horizontal such that $\tan(\theta) = 2y/A^2$. The reflected beam makes an angle with the horizontal which is exactly twice this angle. From the result for $\tan(2x)$ in terms of $\tan(x)$ (see Example 3) we thus know that the reflected beam has a gradient m such that

$$m = \tan(\pi - 2\theta) = -\tan(2\theta) = -2(2y/A^2)/(1 - 4y^2/A^4).$$

This looks like a complicated result but it turns out that the end result is quite simple. To find the x value x_0 at which the beam will cross the axis we note that the beam starts off on the parabola at height y and thus has the initial x coordinate $x = y^2/A^2$. In falling a distance y it will travel a further horizontal distance y/m, where m is the gradient of the line along which it travels. Combining all these facts leads to a result for the x value at which the beam crosses the axis:

$$x_0 = y^2/A^2 - y/m = y^2/A^2 + y(1 - 4y^2/A^4)A^2/(4y)$$

Working out the terms shows that the y^2/A^2 is cancelled out, leaving the simple result $x_0 = A^2/4$. The calculation above was carried out in the plane i.e. in two dimensions. A three dimensional parabolic mirror would be constructed by rotating the parabola around the x axis. Our result shows that all rays arriving parallel to the axis will arrive at the focal point at $x = A^2/4$. If a light source is placed at the focus then a strong parallel beam of light will emerge from the mirror (as in a searchlight).

Example 5

A calculation involving a circle

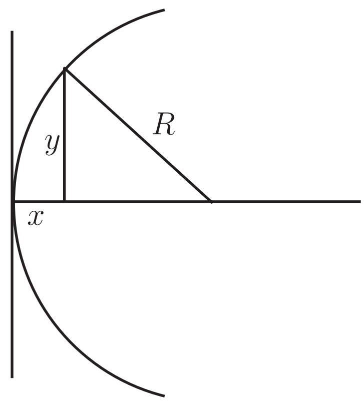
Suppose that the parabolic mirror surface $y = A\sqrt{x}$ of Example 4 is too difficult to make and so is replaced by an arc of a circle. (The full three dimensional parabolic mirror is thus replaced by a portion of a sphere.)

If the circle has radius R, show that it will simulate the parabolic mirror fairly well for light beams striking the mirror close to the x axis and find the x coordinate of the focal point in terms of the radius R.

HINT: Do not consider the reflection of a light beam; try to write the equation of the arc of the circle in the same form as that of the parabola and thus read off the value of $A^2/4$ directly.

Solution

A diagram will clarify the geometry of the situation:



Using the theorem of Pythagoras we see that

$$R^2 = y^2 + (R - x)^2 = R^2 + y^2 + x^2 - 2Rx$$

so that

$$y^2 = x(2R - x)$$

For light beams near to the axis, with x very small, the arc of the circle will be well represented by the equation $y = \sqrt{2Rx}$.

This is identical with the equation of the parabola if we set $A = \sqrt{2R}$.

Thus the focal point of the simulating spherical surface is at the distance $x = A^2/4 = 2R/4 = R/2$ from the centre of the mirror at $x = 0$.

Example 6

Maclaurin series and geometric series

In HELM 16.5 it is shown how to use the geometric series for $1/(1+x)$ to find the Maclaurin series for $\ln(1+x)$ by using the fact that the derivative of $\ln(1+x)$ is $1/(1+x)$. This particular example can be taken further and the same technique can be applied to the $\tan$ function, as illustrated by this problem.

Use the geometric series approach to do the following:

(1) Find the Maclaurin series for $\ln\{(1+x)/(1-x)\}$ and use it to estimate the numerical value of $\ln(2)$.

(2) Find the Maclaurin series for $\tan^{-1}(x)$ and thus obtain an infinite sum of terms which give $\pi/4$.

Solution

(1) We have $\dfrac{d}{dx}\{\ln(1+x)\} = 1/(1+x) = 1 - x + x^2 - x^3 + x^4 \ldots$

Integrating gives

$$\ln(1+x) = x - x^2/2 + x^3/3 - x^4/4 \ldots$$

However, changing sign gives

$$\ln(1-x) = -x - x^2/2 - x^3/3 - x^4/4 \ldots$$

(Note that this makes sense; if a number is less than 1 we expect to get a negative logarithm, since e to a negative index is needed to produce a number less than 1.)

We know that

$$\ln\{(1+x)/(1-x)\} = \ln(1+x) - \ln(1-x).$$

Subtracting the two series in accord with this result gives

$$\ln\{(1+x)/(1-x)\} = 2(x + x^3/3 + x^5/5 \ldots)$$

Note that this series has only odd powers in it, because the function being expanded is odd (like $\tan(x)$). Changing x to $-x$ changes the $\ln$ of a number to the $\ln$ of its reciprocal; this simply changes the sign of the logarithm.

Suppose that we want to find $\ln(Y)$ for some number Y. A little algebra shows that the value of x to use in the series above is given by the formula $x = (Y-1)/(Y+1)$ [work it out!]. To find $\ln(2)$ we thus need to use the small number $x = (2-1)/(2+1) = 1/3$ in the series. Adding the first five terms gives the approximate value 0.693145 for $\ln(2)$, whereas the correct value to 6 d.p. is 0.693147. Using more terms would give greater accuracy.

Solution

(2) We know that $\tan^{-1}(x)$ has derivative $1/(1+x^2)$ and so can set

$$\frac{d}{dx}(\tan^{-1}(x)) = 1/(1+x^2) = 1 - x^2 + x^4 - x^6 + x^8 \ldots$$

Integrating gives

$$\tan^{-1}(x) = x - x^3/3 + x^5/5 - x^7/7 \ldots$$

$\pi/4$ is the angle in the central range which has a tangent equal to 1. Thus we set $x = 1$ in the series to obtain

$$\pi/4 = 1 - 1/3 + 1/5 - 1/7 + 1/9 \ldots$$

The successive numerical sums in this alternating series straddle the exact value of $\pi/4$ and thus give upper and lower bounds to it. This series is not very good for estimating π, however, since it converges very slowly. Two more effective series (which arise in the theory of Fourier series) are the following:

$\pi^2/6 =$ sum of terms $1/n^2$ $\quad (n = 1, 2, 3, \ldots)$, i.e. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$

$\pi^4/90 =$ sum of terms $1/n^4$ $\quad (n = 1, 2, 3, \ldots)$, i.e. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$

Example 7

A link between matrices and complex numbers

Consider the family of 2×2 matrices of the form

$$Z(x, y) = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$

Show that the $Z(x, y)$ matrices multiply together exactly in the same way as complex numbers, if we take $Z(x, y)$ to represent the complex number $x + \mathrm{i}y$. Show also that complex number division can be carried out using this matrix model, with the matrix inverse playing the role of the inverse of its associated complex number.

Solution

The product $Z(a, b)Z(c, d)$ as given by the standard row-column matrix multiplication rule is quickly found to be

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bd & ad + bc \\ -ad - bc & ac - bd \end{pmatrix}$$

The results in the product matrix are exactly the real and imaginary parts of the complex number product $(a+\mathrm{i}b)(c+\mathrm{i}d)$. The rule for the inverse of a 2×2 matrix tells us that the inverse of $Z(x, y)$ is given by

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix}^{-1} = \frac{1}{(x^2 + y^2)} \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

Again, this fits to the inverse of the complex number $(x + \mathrm{i}y)$, which is $(x - \mathrm{i}y)/(x^2 + y^2)$. Thus complex number division can be carried out by multiplying $Z(a, b)$ by the inverse of $Z(c, d)$ to yield correctly the real and imaginary parts of the complex number $(a + \mathrm{i}b)/(c + \mathrm{i}d)$. Complex number multiplication is known to be commutative and we can easily verify that the family of matrices $Z(x, y)$ also commute with one another under matrix multiplication. We have what is called by mathematicians an **isomorphism** (identity of form) between the set of complex numbers $x + \mathrm{i}y$ and the set of 2×2 matrices $Z(x, y)$ with real components x and y. You may check for yourself that complex number addition and subtraction are also correctly described by this matrix model. Note that since we have found an infinite family of matrices which commute with one another under multiplication the statement $AB \neq BA$ which sometimes appears in textbooks on matrices is not correct, strictly speaking, since it appears to be a *universal* statement that matrices *never* commute under multiplication!

Example 8

An approximation for the exponential function

Find an expression for exp(x) by comparing the Maclaurin series for exp(x) with the binomial expansion for $(1+\frac{x}{N})^N$.

Solution

If we compare the exponential series for $\exp(x)$ (often written e^x) with the binomial series for $(1+x/N)^N$, where N is a large integer, then we find that the terms of the binomial series become closer and closer to the terms in $\exp(x)$ as N increases. Suppose, for example, that we wish to work out $e = \exp(1)$. The J^{th} term in the exponential series would be just $1/J!$. The J^{th} term in the binomial expansion of $(1 + 1/N)$ would be $(1/N)^J \ N!/[J!(N-J)!]$ as discussed in HELM 16 on Sequences and Series. This binomial term can be written more usefully in the form

$$\frac{N \times (N-1) \times (N-2) \times \ldots (N+1-J)}{N \times N \times N \times \ldots N} \times (1/J!)$$

For any finite J it is clear that the fraction multiplying $(1/J!)$ becomes closer and closer to 1 as N is increased. Thus the binomial expansion fits more and more closely to the exponential expansion. This leads to an equation involving a limit:

$$\exp(x) = \lim_{N\to\infty} \left\{ \left(1+\frac{x}{N}\right)^N \right\}$$

Note

This equation is not only a celebrated one in classical mathematics; it is actually applied in some areas of scientific research to estimate the exponential of x for difficult cases in which x is not just a number but is a matrix or an operator. Example 9 deals only with the case in which x is a number but it introduces an interesting discovery which has recently appeared in the literature of mathematical education:

'A modification to the e limit', Tony Robin, Mathematical Gazette Vol 88; 2004, pp 279-281.

Example 9

An improved approximation for the exponential function

The function $F(k,x) = (1+x/N)^{kx}$ tends to 1 as the value of N increases for fixed values of k and x. Thus if we modify the classical equation for $\exp(x)$ given above by multiplying the term $(1+x/N)^N$ by $F(k,x)$ we shall *still* obtain $\exp(x)$ in the limit $N \to \infty$. We thus have an infinite family of approximations, with the classical one corresponding to the case $k = 0$. Use a calculator or computer to estimate $e = \exp(1)$ by using the N values $8, 16, 32, \ldots$ (a doubling sequence) and compare the results obtained by using $k = 0$ and $k = 1/2$ in the multiplying factor $F(k,x)$.

Solution

Calculating with a simple 8 digit calculator and rounding the results to 6 significant digits (because of the rounding error in taking high powers of a number) we obtain the following results:

N	$k=0$	$k=1/2$
8	2.56578	2.72143
16	2.63793	2.71911
32	2.67699	2.71849
64	2.69734	2.71833
128	2.70772	2.71828

Thus the $k = 1/2$ approximations reach the correct result much more quickly than the classical $k = 0$ approximations do.

Example 10

An example of an inverse hyperbolic function

Find an expression for the inverse $\tanh$ of x, $\tanh^{-1}(x)$, in terms of the $\ln$ (natural logarithm) function by two different methods:

1. by using algebra to solve the equation $x = \tanh(y)$ for y.
2. by using an approach via geometric series, starting from the fact that the derivative of $\tanh(x)$ is $1 - \tanh^2(x)$.

Solution

(1) Recalling that $\tanh(y) = \sinh(y)/\cosh(y)$ and using the temporary symbol Y for e^y, we have the starting equation

$$x = \tanh(y) = (Y - Y^{-1})/(Y + Y^{-1}) = (Y^2 - 1)/(Y^2 + 1)$$

from which we find that $xY^2 + x = Y^2 - 1$ and thus that

$$Y^2 = e^{2y} = (1 + x)/(1 - x), \quad \text{so that} \quad y = (1/2)\ln\{(1 + x)/(1 - x)\}$$

which expresses the inverse $\tanh$ function, $y = \tanh^{-1}(x)$, in terms of natural logs.

(2) By using an approach which is the same as that used for the inverse tangent function in trigonometry, we find that the inverse $\tanh$ function has the derivative $1/(1 - x^2)$. Using the geometric series approach explained previously we can set

$$\frac{d}{dx}(\tanh^{-1}(x)) = 1/(1 - x^2) = 1 + x^2 + x^4 + \ldots$$

so that by integrating we find

$$\tanh^{-1}(x) = x + x^3/3 + x^5/5 + \ldots$$

Comparison with the result of Example 8 shows that this series is exactly half of the series for the function $\ln\{(1 + x)/(1 - x)\}$, which leads us to the same result as that obtained by algebra in the first part of the solution. Arguments based on the use of infinite series for functions are often used in simple calculus but have the limitation that, strictly speaking, they can only be regarded as trustworthy if the series involved actually converge for the range of x values for which the functions are to be used. The geometric series appearing above are only convergent if x is between 1 and -1 (for the case of real x). Thus a series approach is sometimes only a “shortcut” way of getting a result which can be obtained more rigorously by an alternative method. That is the case for the example treated in this problem. However, if the series involved converges for all x (however large) then a series approach is fully legitimate and there are several such pieces of mathematics involving the series for the exponential and the trigonometric functions.

Physics Case Studies

47.3

Introduction

This Section contains a compendium of case studies involving physics (or related topics) as an additional teaching and learning resource to those included in the previous Workbooks. Each case study may involve several mathematical topics; these are clearly stated at the beginning of each case study.

Prerequisites

Before starting this Section you should ...

- have studied the Sections referred to at the beginning of each Case Study

Learning Outcomes

On completion you should be able to ...

- appreciate the application of various mathematical topics to physics and related subjects

Physics Case Studies

Physics Case Study 1

Black body radiation 1

Mathematical Skills

Topic	Workbook
Logarithms and exponentials	[6]
Numerical integration	[31]

Introduction

A common need in engineering thermodynamics is to determine the radiation emitted by a body heated to a particular temperature at all wavelengths or a particular wavelength such as the wavelength of yellow light, blue light or red light. This would be important in designing a lamp for example. The total power per unit area radiated at temperature T (in K) may be denoted by $E(\lambda)$ where λ is the wavelength of the emitted radiation. It is assumed that a perfect absorber and radiator, called a **black body**, will absorb all radiation falling on it and which emits radiation at various wavelengths λ according to the formula

$$E(\lambda) = \frac{C_1}{\lambda^5\left[e^{C_2/(\lambda T)} - 1\right]} \qquad (1)$$

where $E(\lambda)$ measures the energy (in W m^{-2}) emitted at wavelength λ (in m) at temperature T (in K). The values of the constants C_1 and C_2 are 3.742×10^{-16} W m^{-2} and 1.439×10^{-2} m K respectively. This formula is known as Planck's distribution law. Figure 1.1 shows the radiation $E(\lambda)$ as a function of wavelength λ for various values of the temperature T. Note that both scales are plotted logarithmically. In practice, a body at a particular temperature is not a black body and its emissions will be less intense at a particular wavelength than a black body; the power per unit area radiated by a black body gives the ideal upper limit for the amount of energy emitted at a particular wavelength.

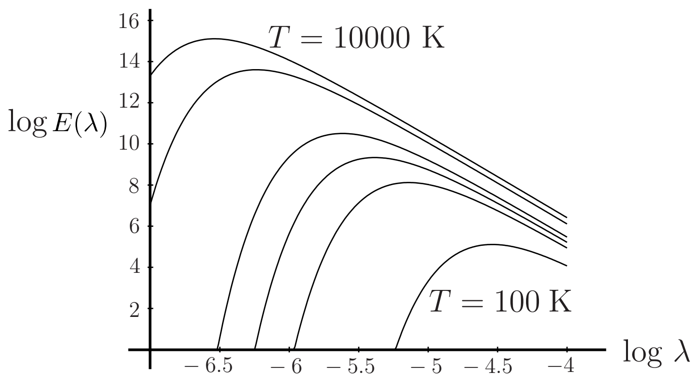


Figure 1.1

The emissive power per unit area $E(\lambda)$ plotted against wavelength (logarithmically) for a black body at temperatures of $T = 100$ K, 400 K, 700 K, 1500 K, 5000 K and 10000 K.

Problem in words

Find the power per unit area emitted for a particular value of the wavelength ($\lambda = 6 \times 10^{-7}$m). Find the temperature of the black body which emits power per unit area ($E(\lambda) = 10^{10}$ W m^{-2}) at a specific wavelength ($\lambda = 4 \times 10^{-7}$m)

Mathematical statement of problem

(a) A black body is at a temperature of 2000 K. Given formula (1), determine the value of $E(\lambda)$ when $\lambda = 6 \times 10^{-7}$m.

(b) What would be the value of T that corresponds to $E(\lambda) = 10^{10}$ W m^{-2} at a wavelength of $\lambda = 4 \times 10^{-7}$m (the wavelength of blue light)?

Mathematical analysis

(a) Here, $\lambda = 6 \times 10^{-7}$ and $T = 2000.$ Putting these values in the formula gives

$$E(\lambda) = 3.742 \times 10^{-16} / \left(6 \times 10^{-7}\right)^5 / \left(e^{1.439\times 10^{-2}/6\times 10^{-7}/2000} - 1\right)$$

$$= 2.98 \times 10^{10} \text{ W m}^{-2} \text{ (to three significant figures).}$$

(b) Equation (1) can be rearranged to give the temperature T as a function of the wavelength λ and the emission $E(\lambda)$.

$$E(\lambda) = \frac{C_1}{\lambda^5\left[e^{C_2/(\lambda T)} - 1\right]} \qquad \text{so} \quad e^{C_2/(\lambda T)} - 1 = \frac{C_1}{\lambda^5 E(\lambda)}$$

and adding 1 to both sides gives

$$e^{C_2/(\lambda T)} = \frac{C_1}{\lambda^5 E(\lambda)} + 1.$$

On taking (natural) logs

$$\frac{C_2}{\lambda T} = \ln\left[\frac{C_1}{\lambda^5 E(\lambda)} + 1\right]$$

which can be re-arranged to give

$$T = \frac{C_2}{\lambda \ln\left[\dfrac{C_1}{\lambda^5 E(\lambda)} + 1\right]} \qquad (2)$$

Equation (2) gives a means of finding the temperature to which a black body must be heated to emit the energy $E(\lambda)$ at wavelength λ.

Here, $E(\lambda) = 10^{10}$ and $\lambda = 4 \times 10^{-7}$ so (2) gives,

$$T = \frac{1.439 \times 10^{-2}}{4 \times 10^{-7} \ln\left[\dfrac{3.742 \times 10^{-16}}{(4 \times 10^{-7}) \times 10^{10}} + 1\right]} = 2380 \text{ K}$$

Interpretation

Since the body is an ideal radiator it will radiate the most possible power per unit area at any given temperature. Consequently any real body would have to be raised to a higher temperature than a black body to obtain the same radiated power per unit area.

Mathematical comment

It is not possible to re-arrange Equation (1) to give λ as a function of $E(\lambda)$ and T. This is due to the way that λ appears twice in the equation i.e. once in a power and once in an exponential. To solve (1) for λ requires numerical techniques but it is possible to use a graphical technique to find the rough value of λ which satisfies (1) for particular values of $E(\lambda)$ and T.

Physics Case Study 2

Black body radiation 2

Mathematical Skills

Topic	Workbook
Logarithms and exponentials	[6]
Numerical solution of equations	[12], [31]

Introduction

A common need in engineering thermodynamics is to determine the radiation emitted by a body heated to a particular temperature at all wavelengths or a particular wavelength such as the wavelength of yellow light, blue light or red light. This would be important in designing a lamp for example. The total power per unit area radiated at temperature T (in K) may be denoted by $E(\lambda)$ where λ is the wavelength of the emitted radiation. It is assumed that a perfect absorber and radiator, called a **black body**, will absorb all radiation falling on it and which emits radiation at various wavelengths λ according to the formula

$$E(\lambda) = \frac{C_1}{\lambda^5\left[e^{C_2/(\lambda T)} - 1\right]} \qquad (1)$$

where $E(\lambda)$ measures the energy (in W m^{-2}) emitted at wavelength λ (in m) at temperature T (in K). The values of the constants C_1 and C_2 are 3.742×10^{-16} W m^{-2} and 1.439×10^{-2} m K respectively. This formula is known as Planck's distribution law. Figure 2.1 shows the radiation $E(\lambda)$ as a function of wavelength λ for various values of the temperature T. Note that both scales are plotted logarithmically. In practice, a body at a particular temperature is not a black body and its emissions will be less intense at a particular wavelength than a black body; the power per unit area radiated by a black body gives the ideal upper limit for the amount of energy emitted at a particular wavelength.

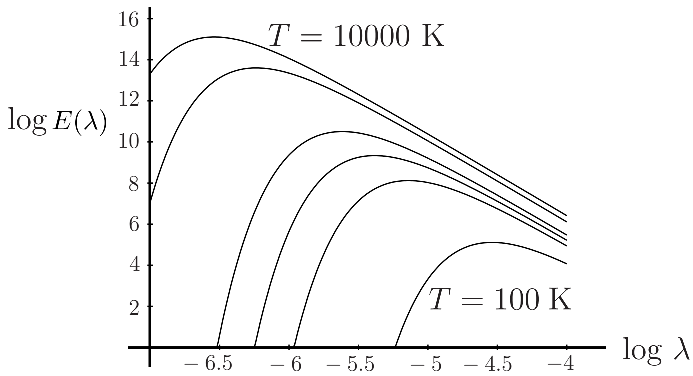


Figure 2.1

The emissive power per unit area $E(\lambda)$ plotted against wavelength (logarithmically) for a black body at temperatures of $T = 100$ K, 400 K, 700 K, 1500 K, 5000 K and 10000 K.

Problem in words

Is it possible to obtain a radiated intensity of 10^8 W m^{-2} at some wavelength for any given temperature?

Mathematical statement of problem

(a) Find possible values of λ when $E(\lambda) = 10^8$ W m^{-2} and $T = 1000$

(b) Find possible values of λ when $E(\lambda) = 10^8$ W m^{-2} and $T = 200$

Mathematical analysis

The graph of Figure 2.2 shows a horizontal line extending at $E(\lambda) = 10^8$ W m^{-2}. This crosses the curve drawn for $T = 1000$K at two points namely once near $\lambda = 10^{-6}$ m and once near $\lambda = 2 \times 10^{-5}$ m. Thus there are two values of λ for which the radiation has intensity $E(\lambda) = 10^8$ W m^{-2} both in the realm of infra-red radiation (although that at $\lambda = 10^{-6}$ m=1 μm is close to the visible light). A more accurate graph will show that the values are close to $\lambda = 9.3 \times 10^{-7}$m and $\lambda = 2.05 \times 10^{-5}$m. It is also possible to use a numerical method such as Newton-Raphson (HELM 12.3 and HELM 31.4) to find these values more accurately. The horizontal line extending at $E(\lambda) = 10^8$ W m^{-2} does not cross the curve for $T = 200$ K. Thus, there is no value of λ for which a body at temperature 200 K emits at $E(\lambda) = 10^8$ W m^{-2}.

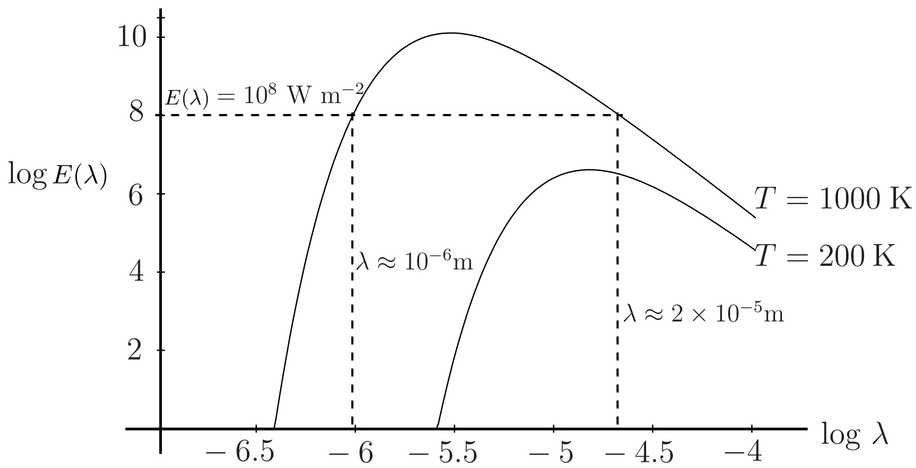


Figure 2.2

The emissive power per unit area $E(\lambda)$ plotted against wavelength (logarithmically) for a black body at temperatures of $T = 200$ K and $T = 1000$ K. For $T = 100$ K an emissive power per unit area of $E(\lambda) = 10^8$ W m^{-2} corresponds to either a wavelength $\lambda \approx 10^{-6}$ or a wavelength $\lambda \approx 2 \times 10^{-5}$. For $T = 200$ K, there is no wavelength λ which gives an emissive power per unit area of $E(\lambda) = 10^8$ W m^{-2}.

Interpretation

Radiation from a black body is dependant both on the temperature and the wavelength. This example shows that it may not be possible for a black body to radiate power at a specific level, irrespective of the wavelength, unless the temperature is high enough.

Physics Case Study 3

Black body radiation 3

Mathematical Skills

Topic	Workbook
Logarithms and exponentials	[6]
Differentiation	[11]

Introduction

A common need in engineering thermodynamics is to determine the radiation emitted by a body heated to a particular temperature at all wavelengths or a particular wavelength such as the wavelength of yellow light, blue light or red light. This would be important in designing a lamp for example. The total power per unit area radiated at temperature T (in K) may be denoted by $E(\lambda)$ where λ is the wavelength of the emitted radiation. It is assumed that a perfect absorber and radiator, called a **black body**, will absorb all radiation falling on it and which emits radiation at various wavelengths λ according to the formula

$$E(\lambda) = \frac{C_1}{\lambda^5 \left[e^{C_2/(\lambda T)} - 1\right]} \qquad (1)$$

where $E(\lambda)$ measures the energy (in W m^{-2}) emitted at wavelength λ (in m) at temperature T (in K). The values of the constants C_1 and C_2 are 3.742×10^{-16} W m^{-2} and 1.439×10^{-2} m K respectively. This formula is known as Planck's distribution law. Figure 3.1 shows the radiation $E(\lambda)$ as a function of wavelength λ for various values of the temperature T. Note that both scales are plotted logarithmically. In practice, a body at a particular temperature is not a black body and its emissions will be less intense at a particular wavelength than a black body; the power per unit area radiated by a black body gives the ideal upper limit for the amount of energy emitted at a particular wavelength.

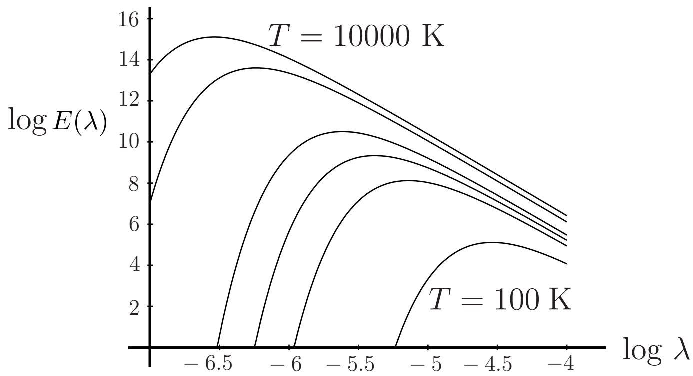


Figure 3.1

The emissive power per unit area $E(\lambda)$ plotted against wavelength (logarithmically) for a black body at temperatures of $T = 100$ K, 400 K, 700 K, 1500 K, 5000 K and 10000 K.

Problem in words

What will be the wavelength at which radiated power per unit area is maximum at any given temperature?

Mathematical statement of problem

For a particular value of T, by means of differentiation, determine the value of λ for which $E(\lambda)$ is a maximum.

Mathematical analysis

Ideally it is desired to maximise

$$E(\lambda) = \frac{C_1}{\lambda^5\left[e^{C_2/(\lambda T)} - 1\right]}.$$

However, as the numerator is a constant, maximising

$$E(\lambda) = \frac{C_1}{\lambda^5\left[e^{C_2/(\lambda T)} - 1\right]}$$

is equivalent to minimising the bottom line i.e.

$$\lambda^5\left[e^{C_2/(\lambda T)} - 1\right].$$

Writing $\lambda^5\left[e^{C_2/(\lambda T)} - 1\right]$ as y, we see that y can be differentiated by the product rule since we can write

$$y = uv \quad \text{where} \quad u = \lambda^5 \quad \text{and} \quad v = e^{C_2/(\lambda T)} - 1$$

so

$$\frac{du}{d\lambda} = 5\lambda^4$$

and

$$\frac{dv}{d\lambda} = -\frac{C_2}{\lambda^2 T}e^{C_2/(\lambda T)} \quad \text{(by the chain rule), Hence}$$

$$\frac{dy}{d\lambda} = \lambda^5\left[-\frac{C_2}{\lambda^2 T}e^{C_2/(\lambda T)}\right] + 5\lambda^4\left[e^{C_2/(\lambda T)} - 1\right]$$

At a maximum/minimum, $\dfrac{dy}{d\lambda} = 0$ hence

$$\lambda^5\left[-\frac{C_2}{\lambda^2 T}e^{C_2/(\lambda T)}\right] + 5\lambda^4\left[e^{C_2/(\lambda T)} - 1\right] = 0$$

i.e.

$$-\frac{C_2}{\lambda T}e^{C_2/(\lambda T)} + 5\left[e^{C_2/(\lambda T)} - 1\right] = 0 \quad \text{(on division by } \lambda^4\text{)}.$$

If we write $C_2/\left(\lambda T\right)$ as z then $-ze^z + 5\left[e^z - 1\right] = 0$ i.e.

$$(5 - z)\,e^z = 5 \qquad (3)$$

This states that there is a definite value of z for which $E(\lambda)$ is a maximum. As $z = C_2/\left(\lambda T\right)$, there is a particular value of λT giving maximum $E(\lambda)$.Thus, the value of λ giving maximum $E(\lambda)$ occurs for a value of T inversely proportional to λ. To find the constant of proportionality, it is necessary to solve Equation (3).

To find a more accurate solution, it is necessary to use a numerical technique, but it can be seen that there is a solution near $z = 5$. For this value of z, e^z is very large ≈ 150 and the left-hand side of (3) can only equal 5 if $5 - z$ is close to zero. On using a numerical technique, it is found that the value of z is close to 4.965 rather than exactly 5.000.

Hence $C_2/(\lambda_{\max} T) = 4.965$ so $\lambda_{\max} = \dfrac{C_2}{4.965T} = \dfrac{0.002898}{T}$.

This relationship is called Wein's law:

$$\lambda_{\max} = \frac{C_w}{T}$$

where $C_w = 0.002898$ m K is known as Wein's constant.

Interpretation

At a given temperature the radiated power per unit area from a black body is dependant only on the wavelength of the radiation. The *nature* of black body radiation indicates that there is a specific value of the wavelength at which the radiation is a maximum. As an example the Sun can be approximated by a black body at a temperature of $T = 5800$ K. We use Wein's law to find the wavelength giving maximum radiation. Here, Wein's law can be written

$$\lambda_{\max} = \frac{0.002898}{5800} \approx 5 \times 10^{-7} \text{ m} = 5000 \text{ Å} \quad \text{(to three significant figures)}$$

which corresponds to visible light in the yellow part of the spectrum.

Physics Case Study 4

Black body radiation 4

Mathematical Skills

Topic	Workbook
Logarithms and Exponentials	[6]
Integration	[13]
Numerical Integration	[31]

Introduction

A common need in engineering thermodynamics is to determine the radiation emitted by a body heated to a particular temperature at all wavelengths or a particular wavelength such as the wavelength of yellow light, blue light or red light. This would be important in designing a lamp for example. The total power per unit area radiated at temperature T (in K) may be denoted by $E(\lambda)$ where λ is the wavelength of the emitted radiation. It is assumed that a perfect absorber and radiator, called a **black body**, will absorb all radiation falling on it and which emits radiation at various wavelengths λ according to the formula

$$E(\lambda) = \frac{C_1}{\lambda^5\left[e^{C_2/(\lambda T)} - 1\right]} \qquad (1)$$

where $E(\lambda)$ measures the energy (in W m^{-2}) emitted at wavelength λ (in m) at temperature T (in K). The values of the constants C_1 and C_2 are 3.742×10^{-16} W m^{-2} and 1.439×10^{-2} m K respectively. This formula is known as Planck's distribution law. Figure 4.1 shows the radiation $E(\lambda)$ as a function of wavelength λ for various values of the temperature T. Note that both scales are plotted logarithmically. In practice, a body at a particular temperature is not a black body and its emissions will be less intense at a particular wavelength than a black body; the power per unit area radiated by a black body gives the ideal upper limit for the amount of energy emitted at a particular wavelength.

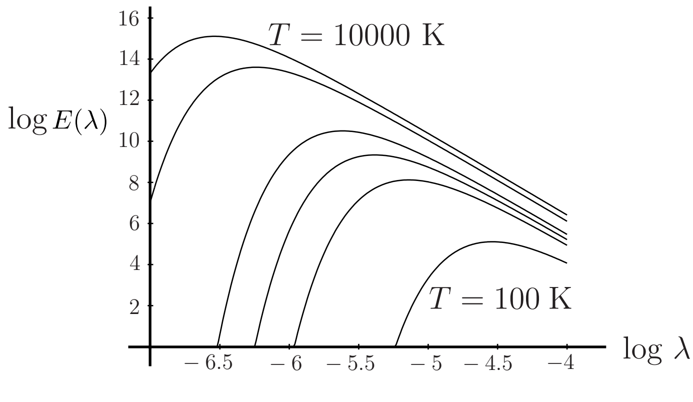


Figure 4.1

The emissive power per unit area $E(\lambda)$ plotted against wavelength (logarithmically) for a black body at temperatures of $T = 100$ K, 400 K, 700 K, 1500 K, 5000 K and 10000 K.

Problem in words

Determine the total power per unit area radiated at all wavelengths by a black body at a given temperature.

The expression (1) gives the amount of radiation at a particular wavelength λ. If this expression is summed across all wavelengths, it will give the total amount of radiation.

Mathematical statement of problem

Calculate

$$E_b = \int_0^\infty E(\lambda)d\lambda = \int_0^\infty \frac{C_1}{\lambda^5\left[e^{C_2/(\lambda T)} - 1\right]}d\lambda.$$

Mathematical analysis

The integration can be achieved by means of the substitution $U = C_2/(\lambda T)$ so that

$$\lambda = C_2/(UT), \quad dU = -\frac{C_2}{\lambda^2 T}d\lambda \quad \text{i.e.} \quad d\lambda = -\frac{\lambda^2 T}{C_2}dU.$$

When $\lambda = 0$, $U = \infty$ and when $\lambda = \infty$, $U = 0$. So E_b becomes

$$E_b = \int_\infty^0 \frac{C_1}{\lambda^5\left(e^U - 1\right)}\left(-\frac{\lambda^2 T}{C_2}\right)dU = -\int_\infty^0 \frac{C_1 T}{C_2\lambda^3\left(e^U - 1\right)}dU$$

$$= \int_0^\infty \frac{C_1 T}{C_2\left(e^U - 1\right)}\left(\frac{UT}{C_2}\right)^3 dU = T^4\int_0^\infty \frac{C_1}{(C_2)^4}\frac{U^3}{\left(e^U - 1\right)}dU.$$

The important thing is that E_b is proportional to T^4 i.e. the total emission from a black body scales as T^4. The constant of proportionality can be found from the remainder of the integral i.e.

$$\frac{C_1}{(C_2)^4}\int_0^\infty \frac{U^3}{\left(e^U - 1\right)}dU \quad \text{where} \quad \int_0^\infty \frac{U^3}{\left(e^U - 1\right)}dU$$

can be shown by means of the polylog function to equal $\dfrac{\pi^4}{15}$. Thus

$$E_b = \frac{C_1\pi^4}{15(C_2)^4}T^4 = 5.67 \times 10^{-8}\ \text{W m}^{-2} \times T^4 = \sigma T^4, \text{ say}$$

i.e.

$$E_b = \sigma T^4$$

This relation is known as the Stefan-Boltzmann law and $\sigma = 5.67 \times 10^{-8}$ W m^{-2} is known as the Stefan-Boltzmann constant.

Interpretation

You will no doubt be familiar with Newton's law of cooling which states that bodies cool (under convection) in proportion to the simple difference in temperature between the body and its surroundings. A more realistic study would incorporate the cooling due to the radiation of energy. The analysis we have just carried out shows that heat loss due to radiation will be proportional to the difference in the fourth powers of temperature between the body and its surroundings.

Physics Case Study 5

Amplitude of a monochromatic optical wave passing through a glass plate

Mathematical Skills

Topic	Workbook
Trigonometric functions	[4]
Complex numbers	[10]
Sum of geometric series	[16]

Introduction

The laws of optical reflection and refraction are, respectively, that the angles of incidence and reflection are equal and that the ratio of the sines of the incident and refracted angles is a constant equal to the ratio of sound speeds in the media of interest. This ratio is the index of refraction (n). Consider a monochromatic (i.e. single frequency) light ray with complex amplitude A propagating in air that impinges on a glass plate of index of refraction n (see Figure 5.1). At the glass plate surface, for example at point O, a fraction of the impinging optical wave energy is transmitted through the glass with complex amplitude defined as At where t is the transmission coefficient which is assumed real for the purposes of this Case Study. The remaining fraction is reflected. Because the speed of light in glass is less than the speed of light in air, during transmission at the surface of the glass, it is refracted toward the normal. The transmitted fraction travels to B where fractions of this fraction are reflected and transmitted again. The fraction transmitted back into the air at B emerges as a wave with complex amplitude $A_1 = At^2$. The fraction reflected at B travels through the glass plate to C with complex amplitude rtA where $r \equiv |r|e^{-i\xi}$ is the complex reflection coefficient of the glass/air interface. This reflected fraction travels to D where a fraction of this fraction is transmitted with complex amplitude $A_2 = A\ t^2r^2e^{-i\varphi}$ where φ is the phase lag due to the optical path length difference with ray 1 (see Engineering Example 4 in HELM 4.2). No absorption is assumed here therefore $|t|^2 + |r|^2 = 1$.

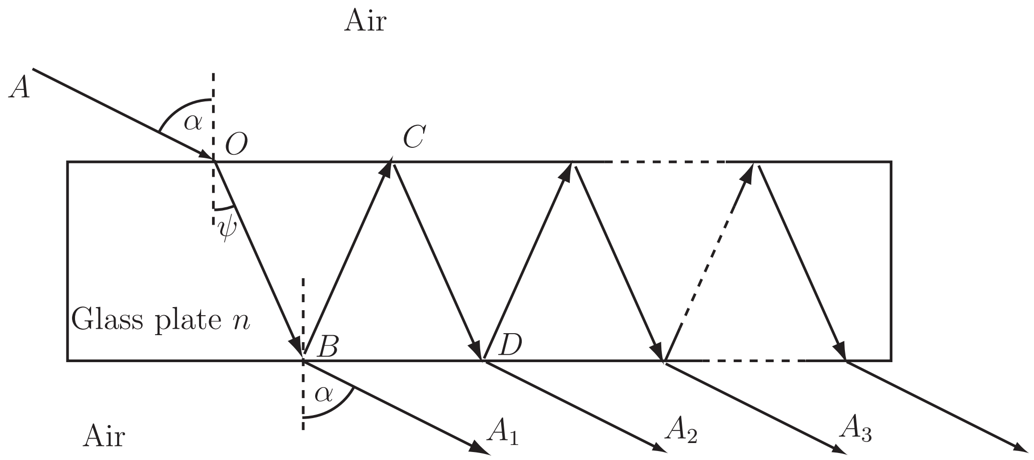


Figure 5.1: Geometry of a light ray transmitted and reflected through a glass plate

Problem in words

Assuming that the internal faces of the glass plate have been treated to improve reflection, and that an infinite number of rays pass through the plate, compute the total amplitude of the optical wave passing through the plate.

Mathematical statement of problem

Compute the total amplitude $\Lambda = \sum_{i=1}^{\infty} A_i$ of the optical wave outgoing from the plate and show that $\Lambda = \dfrac{At^2}{1-|r|^2e^{-i(\varphi+2\xi)}}$. You may assume that the expression for the sum of a geometrical series of real numbers is applicable to complex numbers.

Mathematical analysis

The objective is to find the infinite sum $\Lambda = \sum_{i=1}^{\infty} A_i$ of the amplitudes from the optical rays passing through the plate. The first two terms of the series A_1 and A_2 are given and the following terms involve additional factor $r^2e^{-i\varphi}$. Consequently, the series can be expressed in terms of a general term or rank N as

$$\Lambda = \sum_{i=1}^{\infty} A_i = At^2 + At^2r^2e^{-i\varphi} + At^2r^4e^{-2i\varphi} + \ldots + At^2r^{2N}e^{-iN\varphi} + \ldots \qquad (1)$$

Note that the optical path length difference creating the phase lag φ between two successive light rays is derived in Engineering Example 4 in HELM 4.2. Taking out the common factor of At^2, the infinite sum in Equation (1) can be rearranged to give

$$\Lambda = At^2[1 + \{r^2e^{-i\varphi}\}^1 + \{r^2e^{-i\varphi}\}^2 + \ldots + \{r^2e^{-i\varphi}\}^N + \ldots]. \qquad (2)$$

The infinite series Equation (2) can be expressed as an infinite geometric series

$$\Lambda = At^2 \lim_{n\to\infty} [1 + q + q^2 + \ldots + q^N + \ldots]. \qquad (3)$$

where $q \equiv r^2e^{-i\varphi}$. Recalling from HELM 16.1 that for q real $[1 + q + q^2 + \ldots + q^N + \ldots] = \dfrac{1-q^N}{1-q}$ for $q \neq 1$, we will use the extension of this result to complex q. We verify that the condition $q \neq 1$ is met in this case. Starting from the definition of $q \equiv r^2e^{-i\varphi}$ we write $|q| = |r^2e^{-i\varphi}| = |r^2||e^{-i\varphi}| = |r^2|$. Using the definition $r \equiv |r|e^{-i\varphi}, |r^2| = ||r|^2e^{-2i\xi}| = |r|^2|e^{-2i\xi}| = |r|^2$ and therefore $|q| = |r|^2 = 1 - |t|^2$. This is less than 1 because the plate interior surface is not perfectly reflecting. Consequently, $|q| < 1$ i.e. $q \neq 1$. Equation (3) can be expressed as

$$\Lambda = At^2 \lim_{n\to\infty} \left\{ \frac{1-q^N}{1-q} \right\}. \qquad (4)$$

As done for series of real numbers when $|q| < 1,\ \lim_{N\to\infty} q^N = 0$ and Equation (4) becomes

$$\Lambda = At^2 \frac{1}{1-r^2e^{-i\varphi}}. \qquad (5)$$

Using the definition of the complex reflection coefficient $r \equiv |r|e^{-i\xi}$, Equation (5) gives the final result

$$\Lambda = \frac{At^2}{1-|r|^2e^{-i(\varphi+2\xi)}} \qquad (6)$$

Interpretation

Equation (6) is a complex expression for the amplitude of the transmitted monochromatic light. Although complex quantities are convenient for mathematical modelling of optical (and other) waves, they cannot be measured by instruments or perceived by the human eye. What can be observed is the intensity defined by the square of the modulus of the complex amplitude.

Physics Case Study 6

Intensity of the interference field due to a glass plate

Mathematical Skills

Topic	Workbook
Trigonometric functions	[4]
Complex numbers	[10]

Introduction

A monochromatic light with complex amplitude A propagates in air before impinging on a glass plate (see Figure 6.1). In Physics Case Study 5, the total complex amplitude $\Lambda = \sum_{i=1}^{\infty} A_i$ of the optical wave outgoing from the glass plate was derived as

$$\Lambda = \frac{At^2}{1 - |r|^2 e^{-i(\varphi + 2\xi)}}$$

where t is the complex transmission coefficient and $r \equiv |r|e^{-i\xi}$ is the complex reflection coefficient. ξ is the phase lag due to the internal reflections and φ is the phase lag due to the optical path length difference between two consecutive rays. Note that the intensity of the wave is defined as the square of the modulus of the complex amplitude.

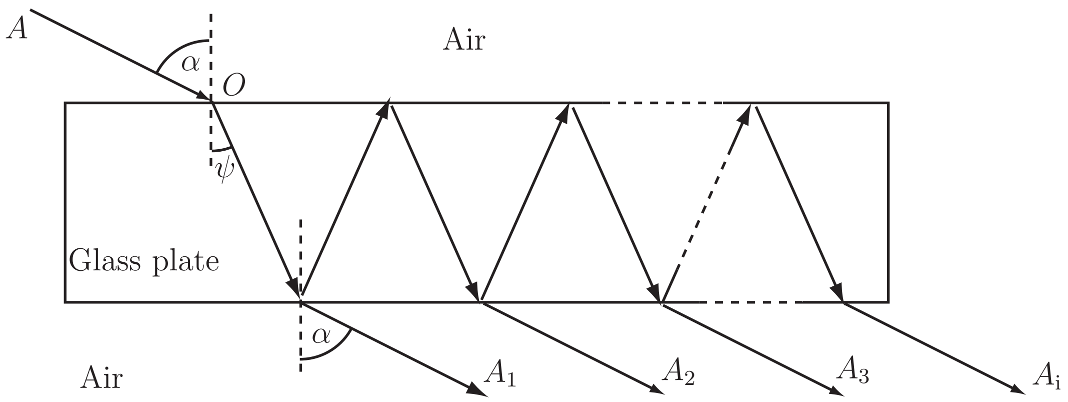


Figure 6.1: Geometry of a light ray transmitted and reflected through a glass plate

Problem in words

Find how the intensity of light passing through a glass plate depends on the phase lags introduced by the plate and the transmission and reflection coefficients of the plate.

Mathematical statement of problem

Defining I as the intensity of the wave, the goal of the exercise is to evaluate the square of the modulus of the complex amplitude expressed as $I \equiv \Lambda\Lambda^*$.

Mathematical analysis

The total amplitude of the optical wave transmitted through the glass plate is given by

$$\Lambda = \frac{At^2}{1 - |r|^2 e^{-i(\varphi+2\xi)}}. \qquad (1)$$

Using the properties of the complex conjugate of products and ratios of complex numbers (HELM 10.1) the conjugate of (1) may be expressed as

$$\Lambda^* = \frac{A^* t^{2^*}}{1 - |r|^2 e^{+i(\varphi+2\xi)}}. \qquad (2)$$

The intensity becomes

$$I \equiv \Lambda\Lambda^* = \frac{AA^* t^2 t^{2^*}}{(1 - |r|^2 e^{-i(\varphi+2\xi)})(1 - |r|^2 e^{+i(\varphi+2\xi)})}. \qquad (3)$$

In HELM 10.1 it is stated that the square of the modulus of a complex number z can be expressed as $|z|^2 = zz^*$. So Equation (3) becomes

$$I = \frac{|A|^2 (tt^*)^2}{(1 - |r|^2 e^{+i(\varphi+2\xi)} - |r|^2 e^{-i(\varphi+2\xi)} + |r|^4)}. \qquad (4)$$

Taking out the common factor in the last two terms of the denominator,

$$I = \frac{|A|^2 |t|^4}{1 + |r|^4 - |r|^2 \{e^{+i(\varphi+2\xi)} + e^{-i(\varphi+2'\xi)}\}}. \qquad (5)$$

Using the exponential form of the cosine function $\cos(\varphi+2\xi) = \{e^{-i(\varphi+2\xi)} + e^{-i(\varphi+2\xi)}\}/2$ as presented in HELM 10.3, Equation (5) leads to the final result

$$I = \frac{|A^2|\ |t|^4}{1 + |r|^4 - 2|r|^2 \cos(2\xi + \varphi)}. \qquad (6)$$

Interpretation

Recall from Engineering Example 4 in HELM 4.3 that φ depends on the angle of incidence and the refractive index of the plate. So the transmitted light intensity depends on angle. The variation of intensity with angle can be detected. A vertical screen placed beyond the glass plate will show a series of interference fringes.

Physics Case Study 7

Propagation time difference between two light rays transmitted through a glass plate

Mathematical Skills

Topic	Workbook
Trigonometric functions	[4]

Introduction

The laws of optical reflection and refraction are, respectively, that the angles of incidence and reflection are equal and that the ratio of the sines of the incident and refracted angles is a constant equal to the ratio of sound speeds in the media of interest. This ratio is the index of refraction (n). Consider a light ray propagating in air that impinges on a glass plate of index of refraction n and of thickness e at an angle α with respect to the normal (see Figure 7.1).

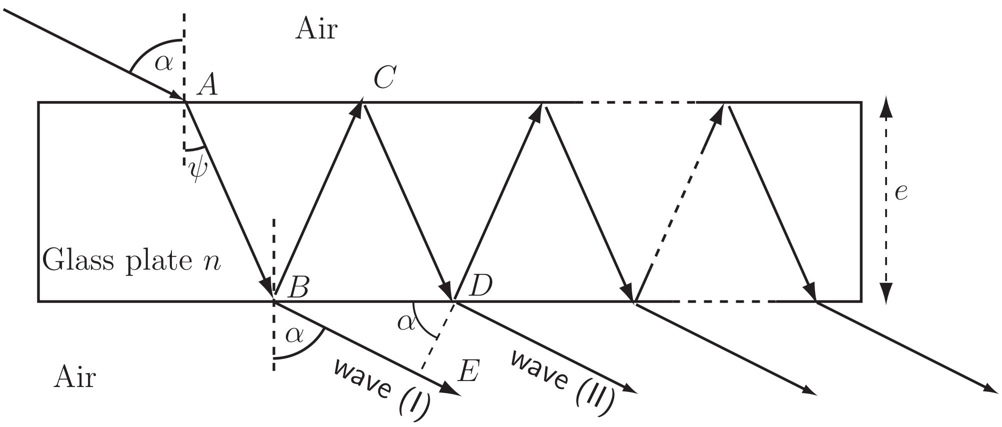


Figure 7.1: Geometry of a light ray transimitted and reflected through a glass plate

At the glass plate surface, for example at point A, a fraction of the impinging optical wave energy is transmitted through the glass and the remaining fraction is reflected. Because the speed of light in glass is less than the speed of light in air, during transmission at the surface of the glass, it is refracted toward the normal at an angle ψ. The transmitted fraction travels to B where a fraction of this fraction is reflected and transmitted again. The fraction transmitted back into the air at B emerges as wave (I). The fraction reflected at B travels through the glass plate to C where a fraction of this fraction is reflected back into the glass. This reflected fraction travels to D where a fraction of this fraction is transmitted as wave (II). Note that while the ray AB is being reflected inside the glass plate at B and C, the fraction transmitted at B will have travelled the distance BE. Beyond the glass plate, waves (I) and (II) interfere depending upon the phase difference between them. The phase difference depends on the propagation time difference.

Problem in words

Using the laws of optical reflection and refraction, determine the difference in propagation times between waves (I) and (II) in terms of the thickness of the plate, the refracted angle, the speed of light in air and the index of refraction. When interpreting your answer, identify three ray paths that are omitted from Figure 1 and state any assumptions that you have made.

Mathematical statement of problem

Using symbols v and c to represent the speed of light in glass and air respectively, find the propagation time difference τ between waves (I) and (II) from $\tau = (BC + CD)/v - BE/c$ in terms of e, n, c and ψ.

Mathematical analysis

The propagation time difference between waves (I) and (II) is given by

$$\tau = (BC + CD)/v - BE/c. \quad (1)$$

As a result of the law of reflection, the angle between the normal to the plate surface and AB is equal to that between the normal and BC. The same is true of the angles to the normal at C, so BC is equal to CD.

In terms of ψ and e

$$BC = CD = e/\cos\psi, \quad (2)$$

so

$$BC + CD = \frac{2e}{\cos\psi}. \quad (3)$$

The law of refraction (a derivation is given in Engineering Example 2 in HELM 12.2), means that the angle between BE and the normal at B is equal to the incident angle and the transmitted rays at B and D are parallel.

So in the right-angled triangle BED

$$\sin\alpha = BE/BD. \quad (4)$$

Note also that, from the two right-angled halves of isosceles triangle ABC,

$$\tan\psi = BD/2e.$$

Replacing BD by $2e\tan\psi$ in (4) gives

$$BE = 2e\tan\psi\sin\alpha. \quad (5)$$

Using the law of refraction again

$$\sin\alpha = n\sin\psi.$$

So it is possible to rewrite (5) as

$$BE = 2en\tan\psi\sin\psi$$

which simplifies to

$$BE = 2en\sin^2\psi/\cos\psi \quad (6)$$

Using Equations (3) and (6) in (1) gives

$$\tau = \frac{2e}{v\cos\psi} - \frac{2en\sin^2\psi}{c\cos\psi}$$

But the index of refraction $n = c/v$ so

$$\tau = \frac{2ne}{c\cos\psi}(1-\sin^2\psi).$$

Recall from HELM 4.3 that $\cos^2\psi \equiv 1-\sin^2\psi$. Hence Equation (6) leads to the final result:

$$\tau = \frac{2ne}{c}\cos\psi \qquad (7)$$

Interpretation

Ray paths missing from Figure 7.1 include reflected rays at A and D and the transmitted ray at C. The analysis has ignored ray paths relected at the 'sides' of the plate. This is reasonable as long as the plate is much wider and longer than its thickness.

The propagation time difference τ means that there is a phase difference between rays (I) and (II) that can be expressed as $\varphi = \dfrac{2\pi c\tau}{\lambda} = \dfrac{4\pi ne\cos\psi}{\lambda}$ where λ is the wave-length of the monochromatic light. The concepts of phase and phase difference are introduced in HELM 4.5 Applications of Trigonometry to Waves. An additional phase shift ξ is due to the reflection of ray (II) at B and C. It can be shown that the optical wave interference pattern due to the glass plate is governed by the phase lag angle $\varphi + 2\xi$. Note that for a fixed incidence angle α (or ψ as the refraction law gives $\sin\alpha = n\sin\psi$), the phase $\varphi + 2\xi$ is constant.

Physics Case Study 8

Fraunhofer diffraction through an infinitely long slit

Mathematical Skills

Topic	Workbook
Trigonometric functions	[4]
Complex numbers	[10]
Maxima and minima	[12]

Introduction

Diffraction occurs in an isotropic and homogeneous medium when light does not propagate in a straight line. This is the case, for example, when light waves encounter holes or obstacles of size comparable to the optical wavelength. When the optical waves may be considered as plane, which is reasonable at sufficient distances from the source or diffracting object, the phenomenon is known as Fraunhofer diffraction. Such diffraction affects all optical images. Even the best optical instruments never give an image identical to the object. Light rays emitted from the source diffract when passing through an instrument aperture and before reaching the image plane. Fraunhofer diffraction theory predicts that the complex amplitude of a monochromatic light in the image plane is given by the Fourier transform of the aperture transmission function.

Problem in words

Express the far-field intensity of a monochromatic light diffracted through an infinitely long slit-aperture characterised by a uniform transmission function across its width. Give your result in terms of the slit-width and deduce the resulting interference fringe pattern. Deduce the changes in the fringe system as the slit-width is varied.

Mathematical statement of problem

Suppose that $f(x)$ represents the transmission function of the slit aperture where the variable x indicates the spatial dependence of transmission through the aperture on the axis perpendicular to the direction of the infinite dimension of the slit. A one-dimensional function is sufficient as it is assumed that there is no variation along the axis of the infinitely long slit. Fraunhofer Diffraction Theory predicts that the complex optical wave amplitude $F(u)$ in the image plane is given by the Fourier transform of $f(x)$ i.e. $F(u) = F\{f(x)\}$. Since the diffracted light intensity $I(u)$ is given by the square of the modulus of $F(u)$, i.e. $I(u) = |F(u)|^2 = F(u)F(u)^*$, the fringe pattern is obtained by studying the minima and maxima of $I(u)$.

Mathematical analysis

Represent the slit width by $2a$. The complex amplitude $F(u)$ can be obtained as a Fourier transform $F(u) = F\{f(x)\}$ of the transmission function $f(x)$ defined as

$$f(x) = 1 \text{ for } -a \leq x \leq a, \quad (1a)$$

$$f(x) = 0 \text{ for } -\infty < x < -a \text{ and } a < x < \infty. \quad (1b)$$

with $f(x) = 1$ or $f(x) = 0$ indicating maximum and minimum transmission respectively, corresponding to a completely transparent or opaque aperture. The required Fourier transform is that of a rectangular pulse (see Key Point 2 in HELM 24.1). Consequently, the result

$$F(u) = 2a\frac{\sin ua}{ua} \quad (2)$$

can be used. The sinc function, $\sin(ua)/ua$ in (2), is plotted in HELM 24.1 page 8 and reproduced below as Figure 8.1.

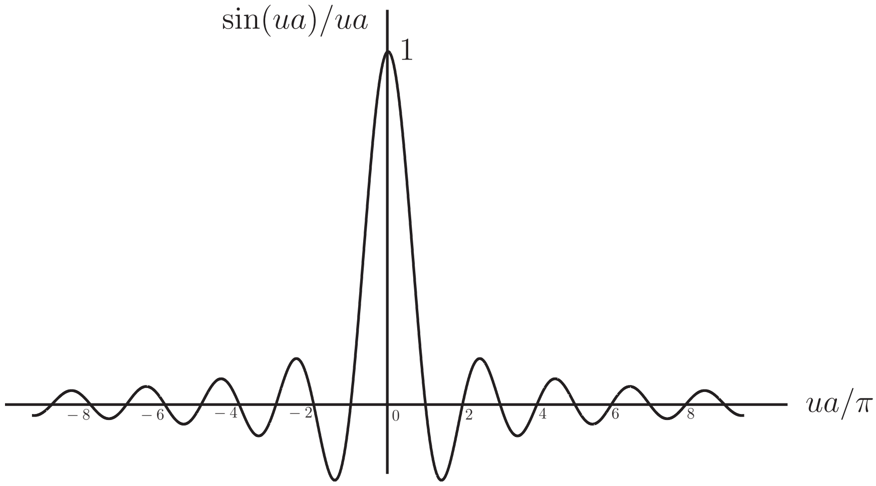


Figure 8.1: Plot of sinc function

$F(u)$ has a maximum value of $2a$ when $u = 0$. Either side of the maximum, Figure 8.1 shows that the sinc curve crosses the horizontal axis at $ua = n\pi$ or $u = n\pi/a$, where n is a positive or negative integer. As u increases, $F(u)$ oscillates about the horizontal axis. Subsequent stationary points, at $ua = (2n+1)\pi/2,\ (|u| \geq \pi/a)$ have successively decreasing amplitudes. Points $ua = 5\pi/2, 9\pi/2 \ldots$ etc., are known as secondary maxima of $F(u)$.

The intensity $I(u)$ is obtained by taking the product of (2) with its complex conjugate. Since $F(u)$ is real, this is equivalent to squaring (2). The definition $I(u) = F(u)F(u)^*$ leads to

$$I(u) = 4a^2\left(\frac{\sin ua}{ua}\right)^2. \quad (3)$$

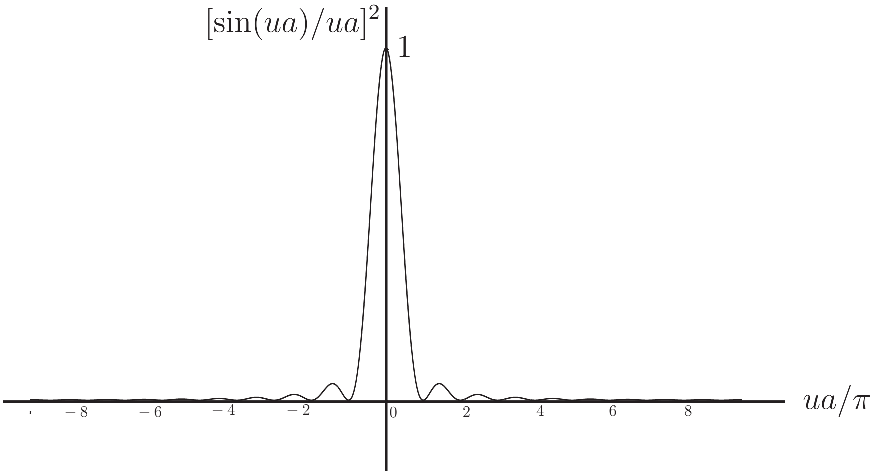


Figure 8.2: Plot of square of sinc function

$I(u)$ differs from the square of the sinc function only by the factor $4a^2$. For a given slit width this is a constant. Figure 8.2, which is a plot of the square of the sinc function, shows that the intensity $I(u)$ is always positive and has a maximum value $I_{\text{max}} = 4a^2$ when $u = 0$. The first intensity minima either side of $u = 0$ occur for $ua = \pm\pi$. Note that the secondary maxima have much smaller amplitudes than that of the central peak.

Interpretation

The transmission function $f(x)$ of the slit aperture depends on the single spatial variable x measured on an axis perpendicular to the direction of the infinite dimension of the slit and no variation of the intensity is predicted along the projection of the axis of the infinitely long slit on the image plane. Consequently, the fringes are parallel straight lines aligned with the projection of the axis of infinite slit length on the image plane (see Figure 8.3). The central fringe at $u = 0$ is bright with a maximum intensity $I_{\text{max}} = 4a^2$ while the next fringe at $u = \pm\pi/a$ is dark with the intensity approaching zero. The subsequent bright fringes (secondary maxima) are much less bright than the central fringe and their brightness decreases with distance from the central fringe.

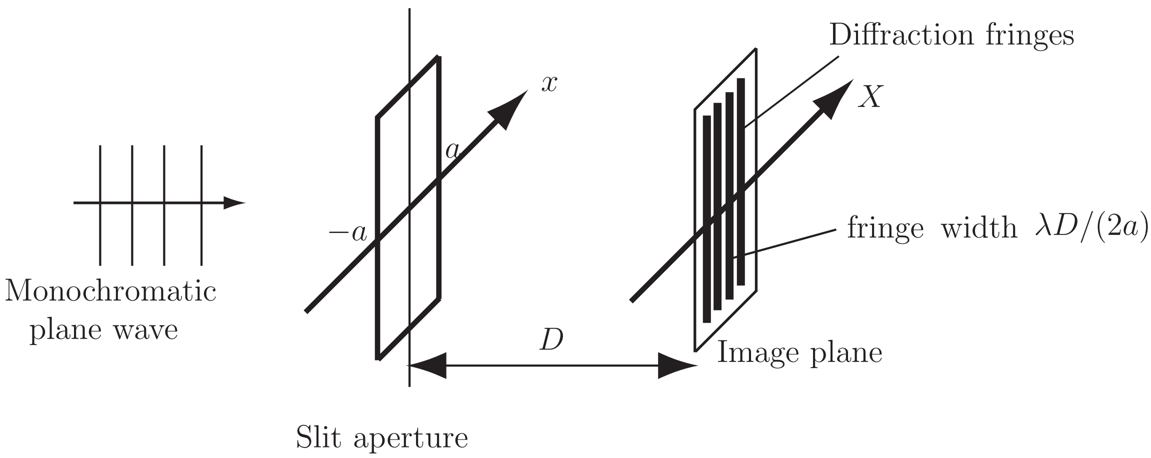


Figure 8.3: Geometry of monochromatic light diffraction through an idealised infinite slit aperture

As the slit-width is increased or decreased, the intensity of the bright central fringe respectively increases or decreases as the square of the slit-width. The Fourier transform variable u is assumed to be proportional to the fringe position X in the image plane. Therefore, as the slit-width a is increased or decreased, the fringe spacing π/a decreases or increases accordingly. It can be shown from diffraction theory that

$$u = \frac{2\pi X}{\lambda D}$$

where λ is the wavelength, D is the distance between the image and aperture planes, and X is the position in the image plane. When $ua = \pm\pi$, $\dfrac{2\pi Xa}{\lambda D} = \pm\pi$ so the first dark fringe positions are given by $X = \pm\,\dfrac{\lambda D}{2a}$. This means that longer wavelengths and longer aperture/image distances will produce wider bright fringes.

Physics Case Study 9

Fraunhofer diffraction through an array of parallel infinitely long slits

Mathematical Skills

Topic	Workbook
Trigonometric functions	[4]
Exponential function	[6]
Complex numbers	[10]
Maxima and minima	[12]
Sum of geometric series	[16]
Fourier transform of a rectangular pulse	[24]
Shift and linearity properties of Fourier transforms	[24]

Introduction

Diffraction occurs in an isotropic and homogeneous medium when light does not propagate in a straight line. This is the case, for example, when light waves encounter holes or obstacles of size comparable to the optical wavelength. When the optical waves may be considered as plane, which is reasonable at sufficient distances from the source or diffracting object, the phenomenon is known as Fraunhofer diffraction. Such diffraction affects all optical images. Even the best optical instruments never give an image identical to the object. Light rays emitted from the source diffract when passing through an instrument aperture and before reaching the image plane. Fraunhofer diffraction theory predicts that at sufficient distance from the diffracting object the complex amplitude of a monochromatic light in the image plane is given by the Fourier transform of the aperture transmission function.

Problem in words

(i) Deduce the light intensity due to a monochromatic light diffracted through an aperture consisting of a single infinitely long slit, characterised by a uniform transmission function across its width, when the slit is shifted in the direction of the slit width.
(ii) Calculate the light intensity resulting from transmission through N parallel periodically spaced infinitely long slits.

Mathematical statement of problem

(i) Suppose that $f(x-l)$ represents the transmission function of the slit aperture where the variable x indicates the spatial dependence of the aperture's transparency on an axis perpendicular to the direction of the infinite dimension of the slit and l is the distance by which the slit is shifted in the negative x-direction. A one-dimensional function is appropriate as it is assumed that there is no variation in the transmission along the length of the slit. The complex optical wave amplitude $G(u)$ in the image plane is give by the Fourier transform of $f(x-l)$ denoted by $G(u) = F\{f(x-l)\}$. The intensity of the diffracted light $I_1(u)$ is given by the square of the modulus of $G(u)$

$$I_1(u) = |F\{f(x-l)\}|^2 = |G(u)|^2. \qquad (1)$$

(ii) In the image plane, the total complex amplitude of the optical wave generated by N parallel identical infinitely long slits with centre-to-centre spacing l is obtained by summing the amplitudes diffracted by each aperture. The resulting light intensity can be expressed as the square of the modulus of the Fourier transform of the sum of the amplitudes. This is represented mathematically as

$$I_N(u) = |F\{\sum_{n=1}^{N} f(x - nl)\}|^2. \qquad (2)$$

Mathematical analysis

(i) Result of shifting the slit in the direction of the slit width

Assume that the slit width is 2a. The complex optical amplitude in the image plane $G(u)$ can be obtained as a Fourier transform $G(u) = F\{f(x - l)\}$ of the transmission function $f(x - l)$ defined as

$$f(x - l) = 1 \text{ for } -a - l \leq x \leq a - l, \qquad (3a)$$

$$f(x - l) = 0 \text{ for } -\infty < x < -a - l \text{ and } a - l < x < \infty. \qquad (3b)$$

The maximum and minimum transmission correspond to $f(x-l) = 1$ and $f(x-l) = 0$ respectively. Note that the function $f(x - l)$ centred at $x = l$ defined by (3a)-(3b) is identical to the function $f(x)$ centred at the origin but shifted by l in the negative x-direction.

The shift property of the Fourier transform introduced in subsection 2 of HELM 24.2 gives the result

$$F\{f(x - l)\} = e^{-iul}F\{f(x - l)\} = e^{-iul}G(u). \qquad (4)$$

Combining Equations (1) and (4) gives

$$I_1(u) = |e^{-iul}G(u)|^2. \qquad (5)$$

The complex exponential can be expressed in terms of trigonometric functions, so

$$|e^{-iul}|^2 = |\cos(ul) - i\sin(ul)|^2.$$

For any complex variable, $|z|^2 = zz^*$, so

$$|e^{-iul}|^2 = [\cos(ul) - i\sin(ul)][\cos(ul) + i\sin(ul)]$$

$$= \cos^2(ul) - i^2\sin^2(ul).$$

Since $i^2 = -1$,

$$|e^{-iul}|^2 = \cos^2(ul) + \sin^2(ul) = 1.$$

The Fourier transform $G(u)$ is that of a rectangular pulse, as stated in Key Point 2 in subsection 3 of HELM 24.1, so

$$G(u) = 2a\frac{\sin ua}{ua} \qquad (6)$$

Consequently, the light intensity

$$I_1(u) = 4a^2\left(\frac{\sin ua}{ua}\right)^2. \qquad (7)$$

This is the same result as that obtained for diffraction by a slit centered at $x = 0$.

Interpretation

No matter where the slit is placed in the plane parallel to the image plane, the same fringe system is obtained.

(ii) Series of infinite slits

Consider an array of N parallel infinitely long slits arranged periodically with centre-to-centre spacing l. The resulting intensity is given by Equation (2). The linearity property of the Fourier transform (see subsection 1 of HELM 24.2) means that

$$F\{\sum_{n=1}^{N} f(x-nl)\} = \sum_{n=1}^{N} F\{f(x-nl)\} \tag{8}$$

Using Equation (4) in (8) leads to

$$F\{\sum_{n=1}^{N} f(x-nl)\} = \sum_{n=1}^{N} e^{-iunl} G(u). \tag{9}$$

The function $G(u)$ is independent of the index n, therefore it can be taken out of the sum to give

$$F\{\sum_{n=1}^{N} f(x-nl)\} = G(u) \sum_{n=1}^{N} e^{-iunl}. \tag{10}$$

Taking the common factor e^{-iul} out of the sum leads to

$$\sum_{n=l}^{N} e^{-iunl} = e^{-iul}\{1 + e^{-iul} + e^{-iu2l} + \ldots e^{-iu(N-1)l}\}. \tag{11}$$

The term in brackets in (11) is a geometric series whose sum is well known (see HELM 16.1).

Assuming that the summation formula applies to complex numbers

$$\sum_{n=l}^{N} e^{-iunl} = e^{-iul} \frac{1 - [e^{-iul}]^N}{1 - e^{-iul}}. \tag{12}$$

Using (12) and (10) in (2) gives an expression for the light intensity

$$I_N(u) = \left| G(u) e^{-iul} \frac{1 - e^{-iuNl}}{1 - e^{-iul}} \right|^2. \tag{13}$$

Recalling that the modulus of a product is the same as the product of the moduli, (13) becomes

$$I_N(u) = |G(u)|^2 \left| e^{-iul} \right|^2 \left| \frac{1 - e^{-iuNl}}{1 - e^{-iul}} \right|^2. \tag{14}$$

Using $|e^{-iul}|^2 = 1$ in (14) leads to

$$I_N(u) = I_1(u) \left| \frac{1 - e^{-iuNl}}{1 - e^{-iul}} \right|^2. \tag{15}$$

The modulus of the ratio of exponentials can be expressed as a product of the ratio and its conjugate which gives

$$\left|\frac{1-e^{-iuNl}}{1-e^{-iul}}\right|^2 = \frac{(1-e^{-iuNl})(1-e^{iuNl})}{(1-e^{-iul})(1-e^{iul})} = \frac{(2-e^{iuNl}-e^{-iuNl})}{2-e^{iul}-e^{-iul}}.$$

Using the definition of cosine in terms of exponentials (see HELM 10.3),

$$\left|\frac{1-e^{-iuNl}}{1-e^{-iul}}\right|^2 = \frac{1-\cos(uNl)}{1-\cos(ul)}.$$

Using the identity $\qquad 1-\cos(2\theta) \equiv 2\sin^2\theta$ gives

$$\left|\frac{1-e^{-iuNl}}{1-e^{-iul}}\right|^2 = \frac{\sin^2\left(\dfrac{uNl}{2}\right)}{\sin^2\left(\dfrac{ul}{2}\right)}. \qquad (16)$$

Using (16) and (5) in (15) leads to the final result for the intensity of the monochromatic light diffracted through a series of N parallel infinitely long periodically spaced slits:

$$I_N(u) = 4a^2\left(\frac{\sin ua}{ua}\right)^2\left(\frac{\sin\left(\dfrac{uNl}{2}\right)}{\sin\left(\dfrac{ul}{2}\right)}\right)^2. \qquad (17)$$

Interpretation

The transmission function $f(x)$ of a single slit depends on the single spatial variable x measured on an axis perpendicular to the direction of the infinite dimension of the slit. The linearity and shift properties of the Fourier transform show that a one-dimensional intensity function of diffracted light is obtained with N identical periodic slits. Consequently, no variation of the intensity is predicted along the projection of the axis of infinite slit length on the image plane. Therefore, the diffraction interference fringes are straight lines parallel to the projection of the axis of the infinite slit on the image plane.

In the expression for the light intensity after diffraction through the N slits, the first term

$$4a^2\left(\frac{\sin ua}{ua}\right)^2$$

is the function corresponding to the intensity due to one slit.

The second factor

$$\left(\frac{\sin\left(\dfrac{Nul}{2}\right)}{\sin\left(\dfrac{ul}{2}\right)}\right)^2$$

represents the result of interference between the waves diffracted through the N slits.

Physics Case Study 10 studies the graphical form of a normalised version of the function in (17) for the case of two slits $(N = 2)$. It is found that the oscillations in intensity, due to the interference term, are bounded by an envelope proportional to the intensity due to one slit.

Physics Case Study 10

Interference fringes due to two parallel infinitely long slits

Mathematical Skills

Topic	Workbook
Trigonometric functions	[4]
Complex numbers	[10]
Maxima and minima	[12]
Maclaurin series expansions	[16]

Introduction

Diffraction occurs in an isotropic and homogeneous medium when light does not propagate in a straight line. This is the case for example, when light waves encounter holes or obstacles of size comparable to the optical wavelength. When the optical waves may be considered as plane, which is reasonable at sufficient distances from the source or diffracting object, the phenomenon is known as **Fraunhofer diffraction**. Such diffraction affects all optical images. Even the best optical instruments never give an image identical to the object. Light rays emitted from the source diffract when passing through an instrument aperture and before reaching the image plane. Prediction of the intensity of monochromatic light diffracted through N parallel periodically spaced slits, idealised as infinite in one direction, is tackled in Physics Case Study 9. The resulting expression for intensity divided by a^2, $2a$ being the slit width, is

$$J_N(u) = 4\left(\frac{\sin ua}{ua}\right)^2 \left(\frac{\sin\left(\frac{uNl}{2}\right)}{\sin\left(\frac{ul}{2}\right)}\right)^2 \qquad (1)$$

where l is the centre-to-centre spacing of the slits and u represents position on an axis perpendicular to that of the infinite length of the slits in the image plane. The first term is called the **sinc** function and corresponds to the intensity due to a single slit (see Physics Case Study 8). The second term represents the result of interference between the N slits.

Problem in words

On the same axes, plot the components (sinc function and interference function) and the normalised intensity along the projection of the slit-width axis on the image plane for a monochromatic light diffracted through two 2 mm wide infinite slits with 4 mm centre-to-centre spacing. Describe the influence of the second component (the interference component) on the intensity function.

Mathematical statement of problem

Plot $y = 4\left(\frac{\sin ua}{ua}\right)^2$, $y = \left(\frac{\sin(ul)}{\sin\left(\frac{ul}{2}\right)}\right)^2$ and $y = J_2(u) = 4\left(\frac{\sin ua}{ua}\right)^2 \times \left(\frac{\sin(ul)}{\sin\left(\frac{ul}{2}\right)}\right)^2$ on the same graph for $a = 1$ mm, $l = 4$ mm and $N = 2$.

[Note that as $\sin(ul) \equiv 2\sin\left(\frac{ul}{2}\right)\cos\left(\frac{ul}{2}\right)$ the expression $\frac{\sin(ul)}{\sin\left(\frac{ul}{2}\right)}$ simplifies to $2\cos\left(\frac{ul}{2}\right)$.]

Mathematical analysis

The dashed line in Figure 10.1 is a plot of the function

$$4\left(\frac{\sin ua}{ua}\right)^2 \qquad (3)$$

The horizontal axis in Figure 10.1 is expressed in units of π/l, $(l = 4a)$, since this enables easier identification of the maxima and minima. The function in (3) involves the square of a ratio of a sine function divided its argument. It has minima (which have zero value, due to the square) when the numerator $\sin(ua) = 0$ and when the denominator $ua \neq 0$. If n is a positive or negative integer, these conditions can be written as $u = n\pi/a$ and $u \neq 0$ respectively. Alternatively, since $l = 4a$, the conditions can be written as $ul/\pi = 4n$ $(n \neq 0)$. This determines the minima of (3) (see Figure 10.1). The first minima are at $n = \pm 1$, i.e. $ul/\pi = \pm 4$. The dashed line shows a maximum at $ul/\pi = 0$ i.e. $u = 0$. When both the sine function and its argument tend to zero $(u = 0)$, the first term in the Maclaurin series expansion of sine (see HELM 16.5) gives the ratio $[(ua)/(ua)]^2 = 1$. So the maximum of (3) at $ul/\pi = 0$ has the value 4. Note that the subsequent maxima of the function (3) are at $ul/\pi = \pm 6$ and the function is not periodic.

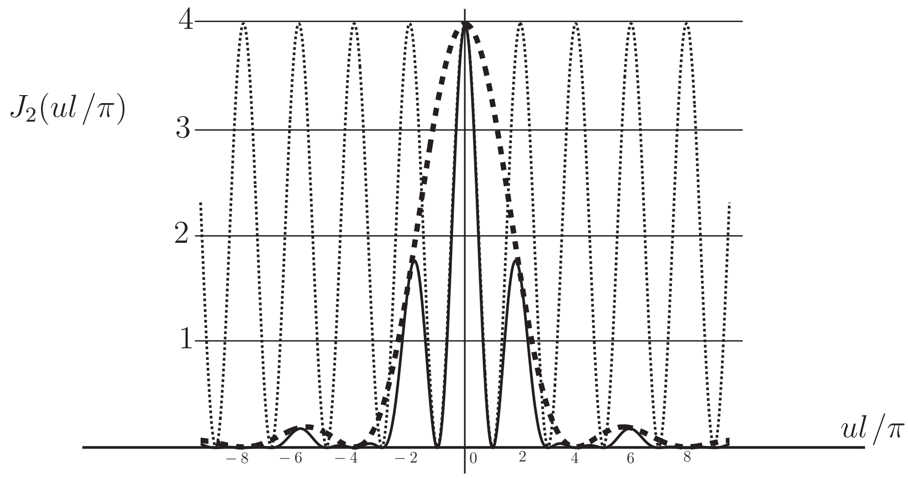


Figure 10.1: Plots of the normalised intensity $J_2(ul/\pi)$ and its component functions

The dotted line is a graph of the interference term

$$\left(\frac{\sin(ul)}{\sin\left(\frac{ul}{2}\right)}\right)^2 \quad \text{which is equivalent to} \quad 4\cos^2\left(\frac{ul}{2}\right) \qquad (4)$$

The interference function (4) is the square of the ratio of two sine functions and has minima (zeros because of the square) when the numerator is $\sin(ul) = 0$ and when the denominator is $\sin(ul/2) \neq 0$. If n is a positive or negative integer, these two conditions can be written as $ul/\pi = n$ and $ul/\pi \neq 2n$. Both conditions are satisfied by the single condition $ul/\pi = (2n+1)$. As a result of scaling the horizontal axis in units of π/l, the zeros of (4) coincide with positive or negative odd integer values on this axis.

When the sine functions in the numerator and denominator of (4) tend to zero simultaneously, the first terms in their Maclaurin series expansions give the ratio $[(ul)/(ul/2)]^2 = 4$ which means that the maxima of (4) have the value 4. They occur when $ul/\pi = 2n$ (see Figure 10.1). Note that the function obtained from the square of the ratio of two functions with periods 2 and 4 (in units of π/l) is a function of period 2.

The solid line is a plot of the product of the functions in (3) and (4) (i.e. Equation (2)) for a slit semi-width $a = 1$ mm and a slit separation with centre-to-centre spacing $l = 4$ mm. For convenience in plotting, the product has been scaled by a factor of 4. Note that the oscillations of the solid line are like those of function (4) but are bound by the dashed line corresponding to the squared sinc function (3). Function (3) is said to provide the **envelope** of the product function (2). The solid line shows a principal maximum at $u = 0$ and secondary maxima around $|u/\pi| \approx 2$ with about half the intensity of the principal maximum. Subsequent higher order maxima show even lower magnitudes not exceeding $1/10^{\text{th}}$ of the principal maximum.

Interpretation

The diffraction interference fringes are parallel straight lines aligned with the projection of the axis of infinite slit length on the image plane as seen in Figure 10.2. The central fringe at $u = 0$ is bright with a maximum normalised intensity $J_2(0) = 4$. Either side of the central fringe at $u/\pi = 1,\ (ul/\pi = 4)$, is dark with the intensity approaching zero. The next bright fringes have roughly half the brightness of the central fringe and are known as secondary. Subsequent bright fringes show even lower brightness. Note that the fringe at $ul/\pi = 6$ is brighter than those at $ul/\pi \approx 3.5$ and $ul/\pi \approx 4.5$ due to the envelope function (3).

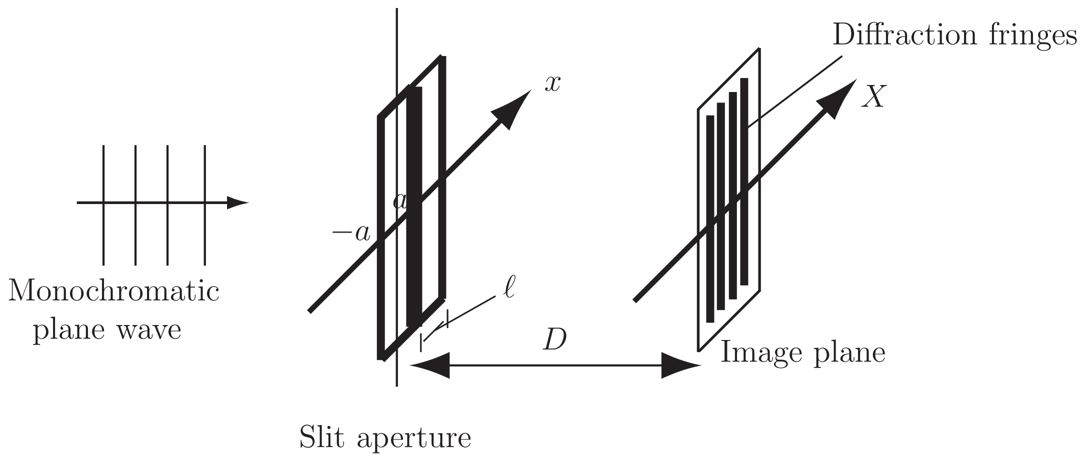


Figure 10.2: Monochromatic light diffraction through two slits

Physics Case Study 11

Acceleration in polar coordinates

Mathematical Skills

Topic	Workbook
Vectors	[9]
Polar coordinates	[17]

Introduction

Consider the general planar motion of a point P whose position is given in polar coordinates. The point P may represent, for example, the centre of mass of a satellite in the gravitational field of a planet.

The position of a point P can be defined by the Cartesian coordinates (x, y) of the position vector $\underline{OP} = x\underline{i} + y\underline{j}$ as shown in Figure 11.1.

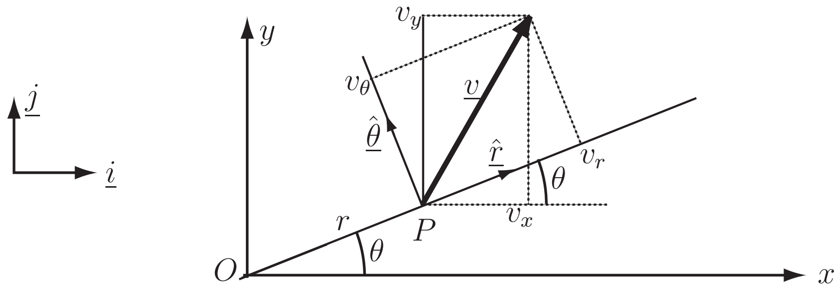


Figure 11.1: Cartesian and polar coordinate system

If $\underline{i}$ and $\underline{j}$ denote the unit vectors along the coordinate axes, by expressing the basis vector set $(\underline{i}, \underline{j})$ in terms of the set $(\hat{\underline{r}}, \hat{\underline{\theta}})$ using trigonometric relations, the components (v_r, v_θ) of velocity $\underline{v} = v_r\hat{\underline{r}} + v_\theta\hat{\underline{\theta}}$ can be related to the components (v_x, v_y) of $\underline{v} = v_x\underline{i} + v_y\underline{j}$ expressed in terms of $(\underline{i}, \underline{j})$.

Problem in words

Express the radial and angular components of the velocity and acceleration in polar coordinates.

Mathematical statement of problem

The Cartesian coordinates can be expressed in terms of the polar coordinates (r, θ) as

$$x = r\cos\theta \qquad (1)$$

and

$$y = r\sin\theta. \qquad (2)$$

The components (v_x, v_y) of velocity $\underline{v} = v_x\underline{i} + v_y\underline{j}$ in the frame $(\underline{i}, \underline{j})$ are derived from the time derivatives of (1) and (2). $\dfrac{d\underline{i}}{dt} = \underline{0}$ and $\dfrac{d\underline{j}}{dt} = \underline{0}$ as the unit vectors along Ox and Oy are fixed with time. The components (a_x, a_y) of acceleration in the frame $(\underline{i}, \underline{j})$ are obtained from the time derivative of $\underline{v} = \dfrac{d}{dt}(\underline{OP})$, and use of expressions for (v_r, v_θ) leads to the components (a_r, a_θ) of acceleration.

Mathematical analysis

The components (v_x, v_y) of velocity $\underline{v}$ are derived from the time derivative of (1) and (2) as

$$v_x = \frac{dx}{dt} = \frac{dr}{dt}\cos\theta - r\frac{d\theta}{dt}\sin\theta \qquad (3)$$

$$v_y = \frac{dy}{dt} = \frac{dr}{dt}\sin\theta - r\frac{d\theta}{dt}\cos\theta. \qquad (4)$$

The components (a_x, a_y) of acceleration are obtained from the time derivatives of (3) and (4),

$$a_x = \frac{d^2x}{dt^2} = \frac{d^2r}{dt^2}\cos\theta - 2\frac{dr}{dt}\frac{d\theta}{dt}\sin\theta - r\left(\frac{d\theta}{dt}\right)^2\cos\theta - r\frac{d^2\theta}{dt^2}\sin\theta \qquad (5)$$

$$a_y = \frac{d^2y}{dt^2} = \frac{d^2r}{dt^2}\sin\theta + 2\frac{dr}{dt}\frac{d\theta}{dt}\cos\theta - r\left(\frac{d\theta}{dt}\right)^2\sin\theta + r\frac{d^2\theta}{dt^2}\cos\theta. \qquad (6)$$

The components (v_x, v_y) of velocity $\underline{v} = v_x\underline{i} + v_y\underline{j}$ and (a_x, a_y) of acceleration $\underline{a} = a_x\underline{i} + a_y\underline{j}$ are expressed in terms of the polar coordinates. Since the velocity vector $\underline{v}$ is the same in both basis sets,

$$v_x\underline{i} + v_y\underline{j} = v_r\underline{\hat{r}} + v_\theta\underline{\hat{\theta}}. \qquad (7)$$

Use of known expresions for the basis vector $(\underline{i}, \underline{j})$ in terms of the basis vectors $(\underline{\hat{r}}, \underline{\hat{\theta}})$ leads to expressions for (v_r, v_θ) in terms of the polar coordinates.

Projection of the basis vectors $(\underline{i}, \underline{j})$ onto the basis vectors $(\underline{\hat{r}}, \underline{\hat{\theta}})$ leads to (see Figure 11.2)

$$\underline{i} = \cos\theta\underline{\hat{r}} - \sin\theta\underline{\hat{\theta}} \qquad (8)$$

$$\underline{j} = \sin\theta\underline{\hat{r}} + \cos\theta\underline{\hat{\theta}} \qquad (9)$$

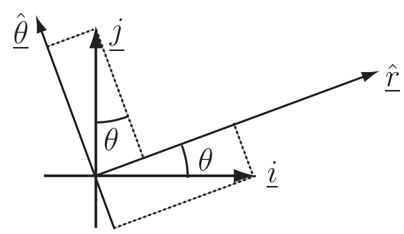


Figure 11.2: Projections of the Cartesian basis set onto the polar basis set

Equation (7) together with (8) and (9) give

$$v_x(\cos\theta\underline{\hat{r}} - \sin\theta\underline{\hat{\theta}}) + v_y(\sin\theta\underline{\hat{r}} + \cos\theta\underline{\hat{\theta}}) = v_r\underline{\hat{r}} + v_\theta\underline{\hat{\theta}}. \qquad (10)$$

The basis vectors $(\underline{\hat{r}}, \underline{\hat{\theta}})$ have been chosen to be independent, therefore (10) leads to the two equations,

$$v_r = v_x \cos\theta + v_y \sin\theta \qquad (11)$$

$$v_\theta = -v_x \sin\theta + v_y \cos\theta. \qquad (12)$$

Using (3) and (4) in (11) and (12) gives

$$v_r = \frac{dr}{dt} \qquad (13)$$

$$v_\theta = r\frac{d\theta}{dt} \qquad (14)$$

Following the same method for the components (a_r, a_θ) of the acceleration,
the equation $\underline{a} = a_x\underline{i} + a_y\underline{j} = a_r\underline{\hat{r}} + a_\theta\underline{\hat{\theta}}$ allows us to write

$$a_r = a_x \cos\theta + a_y \sin\theta \qquad (15)$$

$$a_\theta = -a_x \sin\theta + a_y \cos\theta. \qquad (16)$$

Using (5) and (6) in (15) and (16) leads to

$$a_r = \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 \qquad (17)$$

$$a_\theta = 2\frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}. \qquad (18)$$

Interpretation

The angular velocity $\frac{d\theta}{dt}$ and acceleration $\frac{d^2\theta}{dt^2}$ are often denoted by ω and α respectively. The component of velocity along $\underline{\hat{\theta}}$ is $r\omega$. The component of acceleration along $\underline{\hat{r}}$ includes not only the so-called radial acceleration $\frac{d^2r}{dt^2}$ but also $-r\omega^2$ the centripetal acceleration or the acceleration toward the origin which is the only radial terms that is left in cases of circular motion. The acceleration along $\underline{\hat{\theta}}$ includes not only the term $r\alpha$, but also the Coriolis acceleration $2\frac{dr}{dt}\omega$. These relations are useful when applying Newton's laws in a polar coordinate system. Engineering Case Study 13 in HELM 48 uses this result to derive the differential equation of the motion of a satellite in the gravitational field of a planet.