NOTES ON JUNIOR HIGH SCHOOL PHYSICS

2nd Edition

A.J. Emmerson

FOREWORD

Simplification is no excuse for lessened intellectual rigor. In simplifying science, so that it can be taught to students new to the concepts, we should neither distort nor diminish the underlying principles. Nor should we, in summarizing the facts, present them inaccurately or incompletely.

Falling short in these ways leaves gaps in knowledge and organisation of knowledge that deprive the student of a sound basis for future understanding of science and engineering.

We should not carelessly allow precision in our definition of terms to exceed the precision of our knowledge. We should not imply knowledge when it does not exist.

We should not insist on "simplicity" when what is really required is clarity. Doing so can de-emphasize or distort the connection to broader implications and principles. Over simplification breeds confusion.

Once entrenched, over simplification is difficult to throw out.

It is with that in mind that I have, perhaps presumptuously, prepared these notes.

High school students of an earlier age coped well with a rigorous and detailed science-syllabus. We should not under rate the present generation.

I wonder whether teachers feel that the syllabus is unbalanced. Instructional depth in one field or topic at the expense of another is, of course, always a matter of judgement. Why should a comprehensive review of nuclear power generation displace all discussion of propulsion engines. Is it ever reasonable to present the minutiae of technical data and jargon of one quite specialised topic in the face of a misleadingly perfunctory presentation of a core topic such as standards and units of measurement.

However, that is a debate for another day. In the meantime, we must answer the question, "Why is it so?"

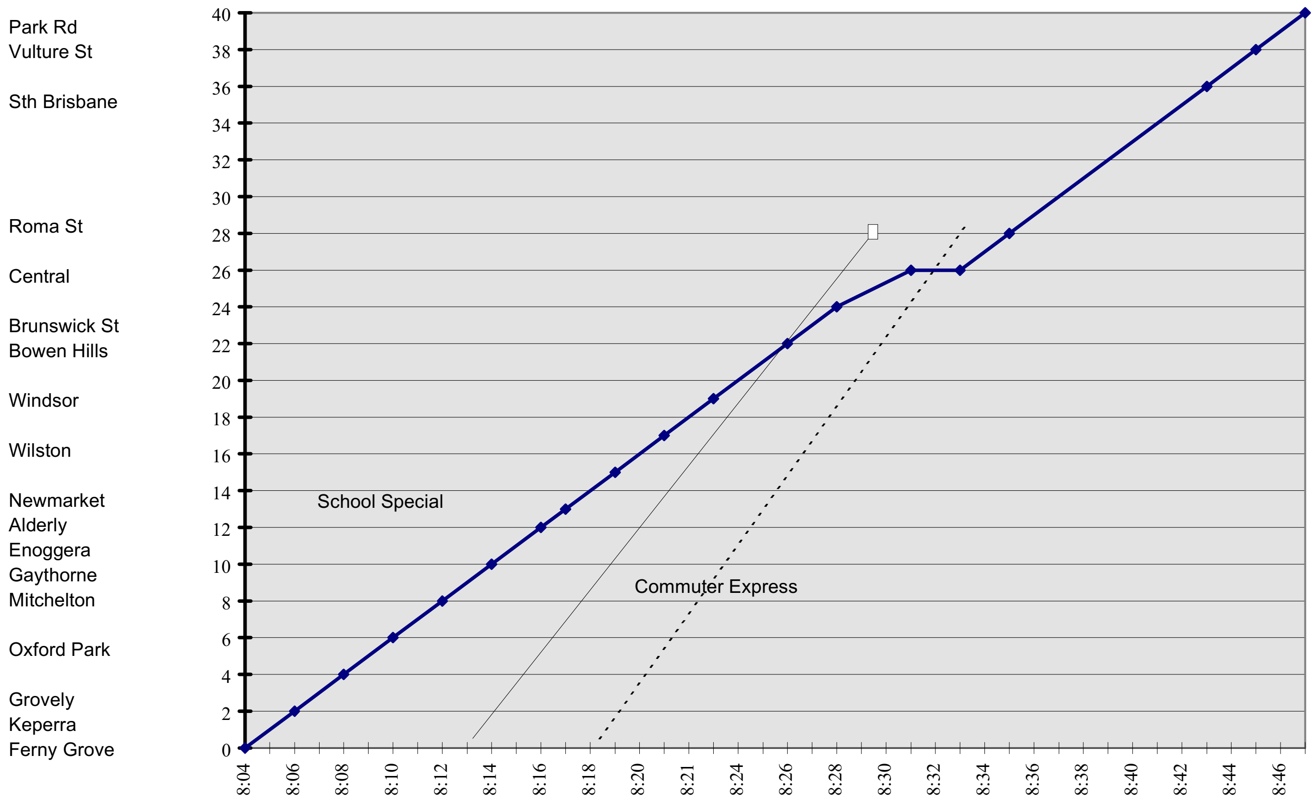
Alan Emmerson
September 2001

CONTENTS

TIME FLIES LIKE AN ARROW - THE GRAPHICAL REPRESENTATION OF SPEED AND ACCELERATION

KEEPING THE TRAINS RUNNING

City Train's School Special leaves Ferny Grove station at 8:04am and travels to Park Road stopping at all stations along the way. It's progress can be seen on the following chart. Notice where the train waits at Central (Platform 3) for most of the passengers to disembark.



City Train marketing manager proposes to introduce a new service, the Commuter Express. He at first proposes that this train should leave Ferny Grove at 8:18am, travel at best average safe speed, and stop only at Roma St after passing through Central (Platform 3). This proposal is shown by the dotted line on the chart.

However, the Train Controller notices that when the Commuter Express arrives at Central, Platform 3 will be occupied by the School Special., and there is certain to be a collision.

The Train Controller can see that if he slides to the left the dotted line representing the journey of the Express, he can eliminate the problem. He moves the line to the position shown by the solid line on the chart. The Express now leaves Ferny Grove at 8:13am and catches up with and passes the School Special on the passing track between Windsor and Bowen Hills

This just one example of how useful it is to be able to draw and interpret speed and distance charts.

SOME DEFINITIONS

Distance and Displacement

The distance travelled by a moving object since it started to move; that is, the distance the object has travelled away from its starting point or origin, is called the *displacement* of the object. Displacement is another word for distance travelled from the beginning.

Displacement is measured in the same units as length, that is, the preferred unit is the metre. The abbreviation for metre is m.

We can use decimal multiples and sub multiples of the metre of course. The preferred multiples and sub multiples are the "ternary" powers of 10 including nano ,micro, milli, and kilo.

Time

When we measure time we are really only measuring the length of a relatively short period of time between specified events. We measure *intervals of time*.

The unit of measurement for intervals of time is the second. The abbreviation for second is s

Again we can use decimal multiples and sub multiples of the second. The preferred multiples and sub multiples are once more the "ternary" powers of 10 including nano, micro, and milli.

Other units for the measurement of intervals of time are acceptable for special purposes. These include the minute and the hour.

Time Zero When observing how something moves or otherwise changes with time it is often convenient to say that the observation began, or some initial event happened, at time zero, that is when the "value" of time was zero.

This leads to the notion that time after zero is a positive value; and *negative values of time correspond to time before time zero.* This is merely a useful mathematical convenience.

Speed

In a motor car we think of speed as how fast we are going, but scientifically speaking, speed is the rate at which our distance from some specified place is increasing. *Speed is increase in displacement per unit of time*.

Speed is measured in the "derived unit" metres per second. The abbreviation for the unit of speed is ms^{-1} or m/s .

We can use decimal multiples and submultiples of the m/s. For example kilo m/s, or milli m/s. The correct way to write this is without a space between the prefix and the unit. So that we should write km/s and mm/s.

A negative value of speed means that the displacement is decreasing.

A WORD ABOUT GRAPHS

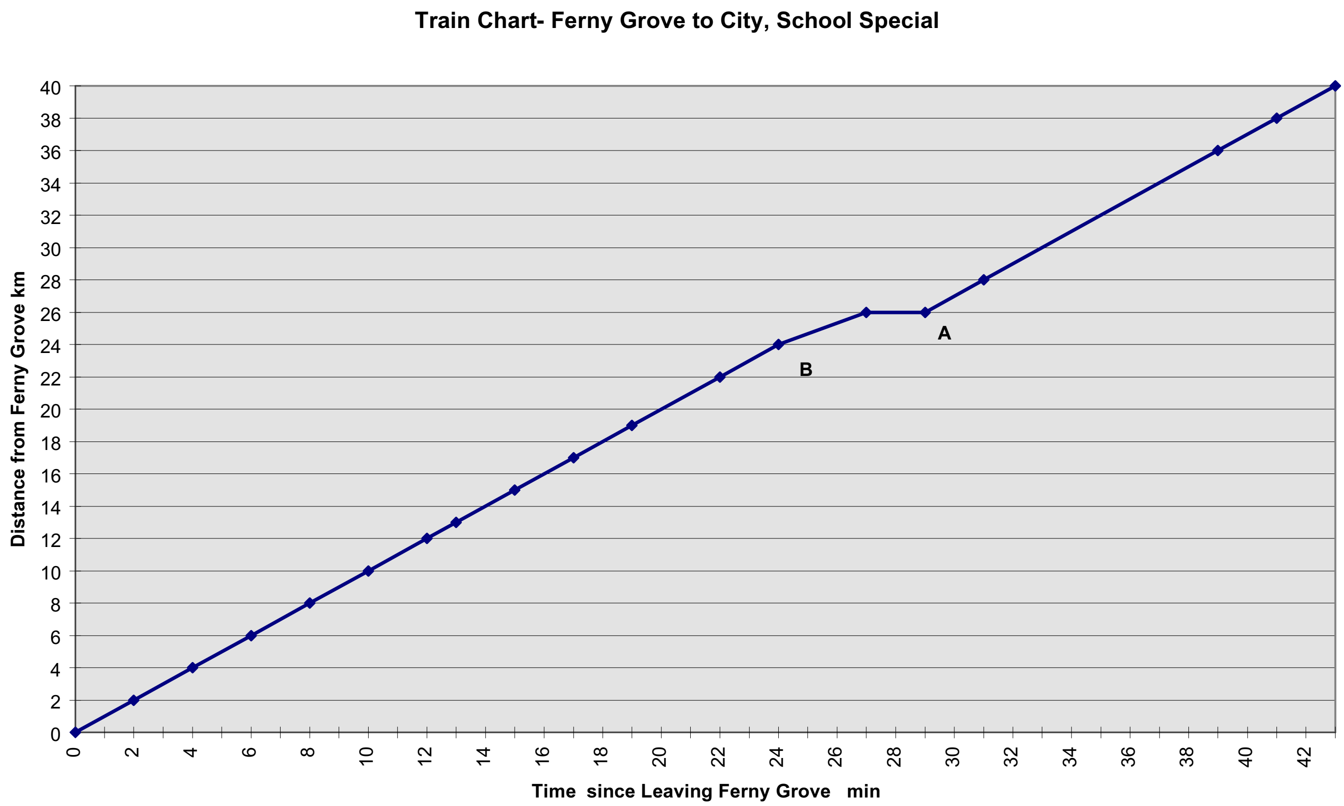
There are some unsatisfactory aspects of the train chart we have just seen. They make it difficult to understand quickly. Firstly, there is no title. All charts or graphs should have a title stating clearly what the chart is illustrating.

Secondly, the axes are not labelled. Each axis on a graph should be clearly labelled with the name of the quantity or "variable" set out along that axis, together with the units in which it is being measured. Each axis should be marked in a scale that permits important changes in the value of the data to be clearly seen. A hand drawn graph smaller than half an A4 page is seldom satisfactory.

The observations or results are called the "dependent variable". *The dependent variable is set out on the vertical axis*.

The horizontal axis should be used for the variable which we control in our experiment or which we think is responsible for the changes observe. We call this the "independent variable". The axis may be marked to show categories or it may be graduated as a continuous number line . Unless you have a very good reason to do otherwise, the axes should meet or cross at the left hand bottom corner with the value zero on both scales..

See what our train chart looks like when we apply those rules.



Speed

We can now see that this is a graph of distance gone (displacement) versus time. We can see that in the first twelve minutes (that is from 8:04 to 8:16) the train has travelled 12km (displacement increases from zero to 12km). The average speed of the train in this time must have been 12km/12min that is 16.7m/s.

What is the average speed over the next 12 minutes ?

On the revised train chart, draw in the journey of the commuter express. What is its average speed from Ferny Grove to Roma St?

From the graph determine the displacement increase between 24 and 27 minutes. In three minutes the train has travelled two kilometres. That is only 11.1 m/s. The train has slowed to go through the tunnel. You will notice that the graph over this portion of the journey is not as steep as elsewhere. The steepness or "slope" of the graph of displacement versus time represents speed. The greater the slope the greater the speed.

Acceleration

At point A on the train chart you can see that the slope of the graph increases. This represents the train gathering speed on leaving the platform. Increasing speed is called accelerating.

Acceleration is the rate at which the speed increases. In other words, acceleration is *the increase in speed per unit time.*

Accordingly, acceleration is measured as the change in the number of metres per second that happens in one second. That is, as metres per second per second. We abbreviate this as m/s/s, or m/s^2 which is mathematically equivalent.

We can use decimal multiples and submultiples of the m/s. For example km/s/s and mm/s/s.

A negative value of acceleration means that the speed is decreasing. The moving object is "decelerating" The School Special can be seen decelerating at point B.

Clever students will have noticed that at points A and B the train appeared to change speed very quickly. - instantaneously in fact. Real trains do not do that of course. If we had measured the displacement of the train at shorter intervals, that artificiality would have partly disappeared.

GRAPHS OF SPEED VS TIME

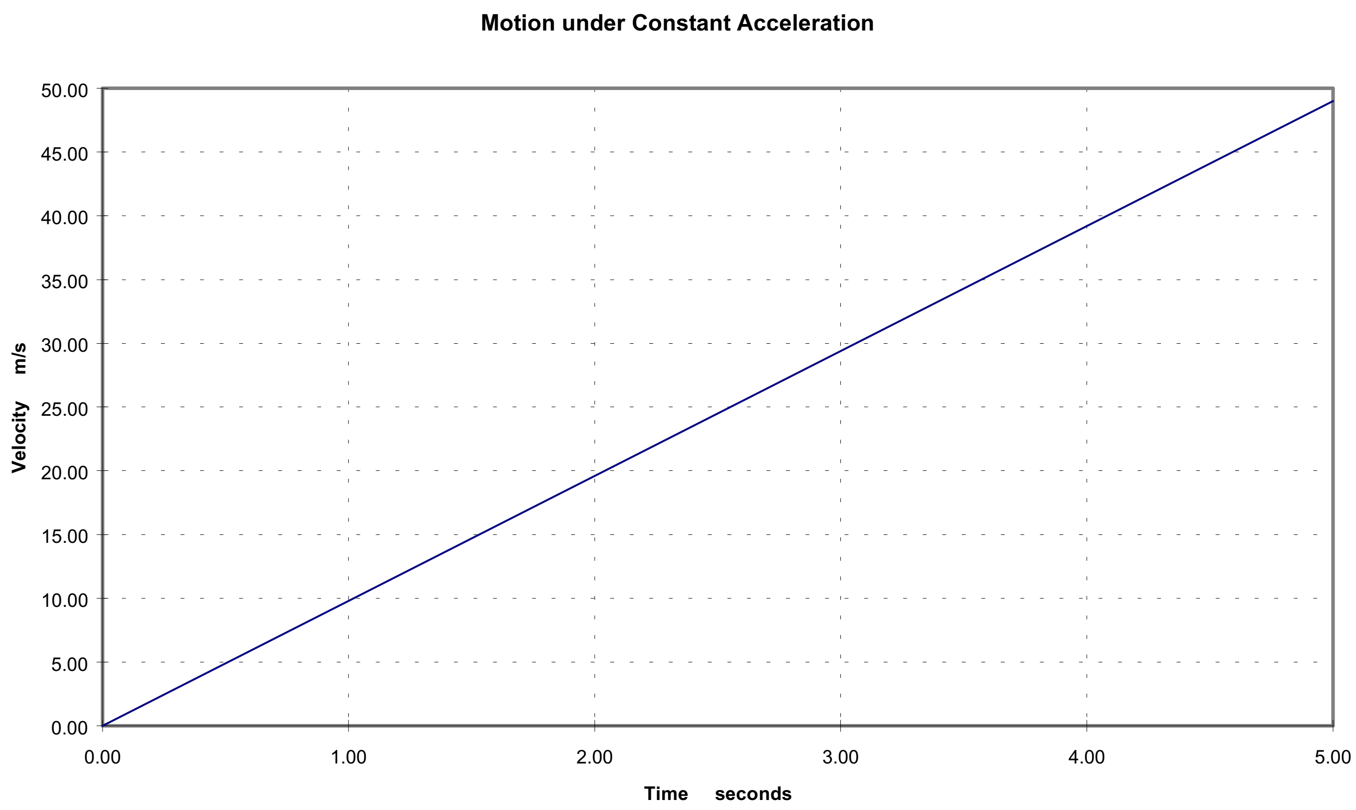
Some times it is convenient to use a graph of speed versus time to describe the movement of an object. In this type of graph, time is represented on the horizontal axis and speed is represented on the vertical axis.

The slope of the graph represents the acceleration of the object - the amount the speed changes in a given time.

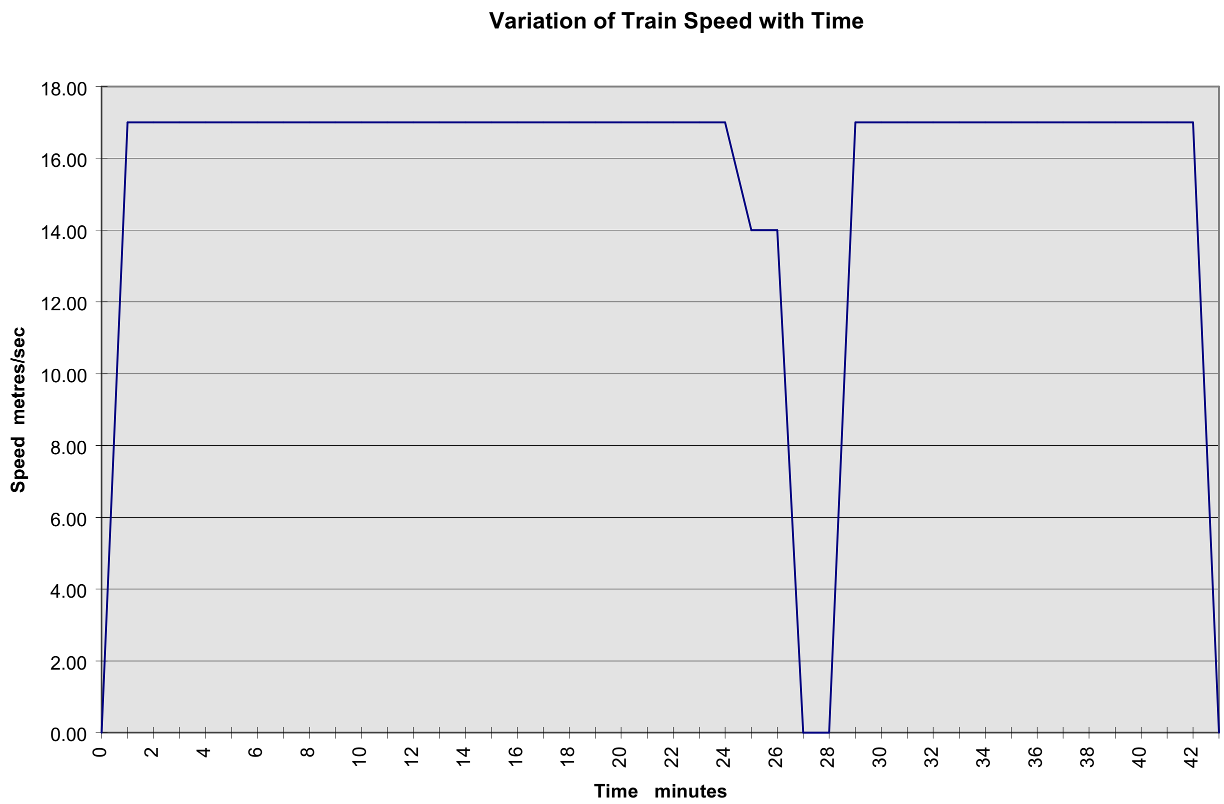
The area under the graph between one time and another represents the displacement of the object between those times.

The following graph represents the motion of an object which experiences a constant acceleration of 9.8metres/second/second for five seconds.

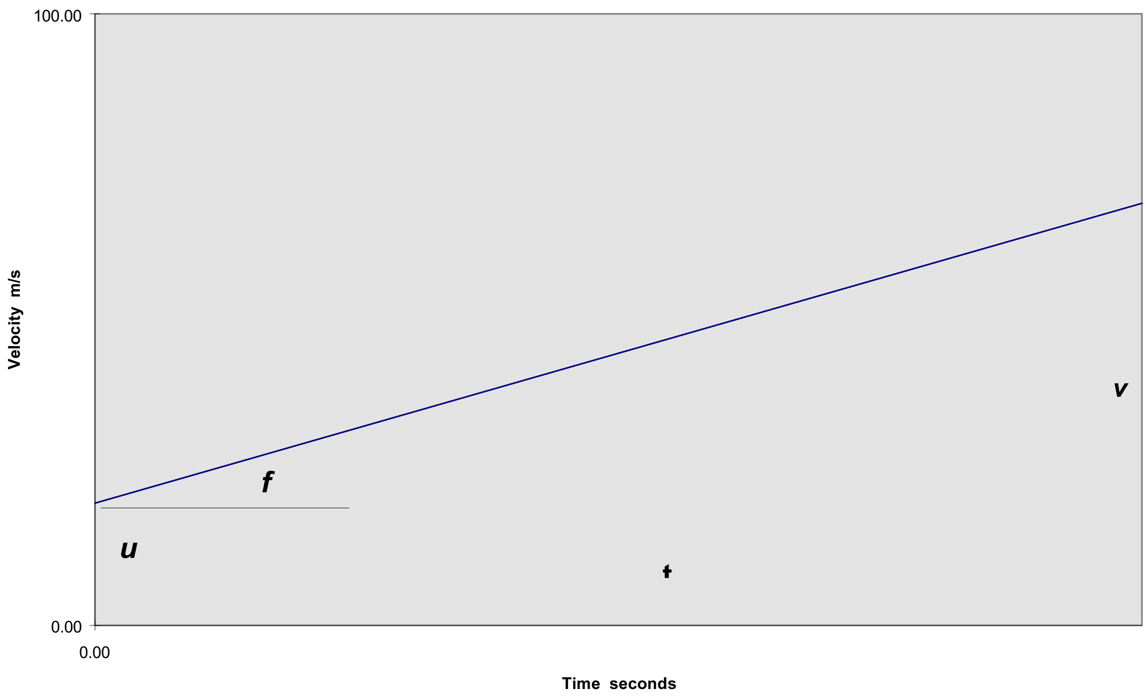
Where do you think you might observe such a motion?



This is what the journey of the Schools Special looks like when plotted on a speed versus time graph.



Your teacher might use a diagram like this to derive the following formulae which you will find very useful.



$$v = u + ft$$

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

$$v^2 = u^2 + 2fs$$

Where:

u is the initial velocity that is the velocity at time zero,

v is the final velocity that is the velocity at time t

s is the displacement or distance travelled between time zero and time t

ENERGY

WORK ENERGY AND POWER

When we push on something and move it, scientists and engineers say we are doing "mechanical work" or "work" for short. Machines which push or pull things and move them are also doing work.

To do work you need energy. Energy is defined as the ability to do work..

Things have energy as a result of :

(a) their mass and speed, - "kinetic energy",

(b) their height above the earth - gravitational potential energy (or just "potential energy”),

(kinetic and potential energy are often considered together and called "mechanical energy")

(c) their possibility of having chemical reactions - chemical potential energy (or just "chemical energy”),

(d) their temperature - the energy which results from being hot is called "heat",

(e) their being stretched - "strain energy",

(f) the forces which hold their atoms together - "atomic" or "nuclear energy",

Devices which store electricity can have "electric energy". Many substances are made of groups of atoms, called molecules. These molecules vibrate and the associated energy is called "internal energy" Gases have energy due to their pressure and volume as well as their temperature. This is called "enthalpy.".

Units of Measurement

Just as we measure distances in metres and time in seconds, so we have units of measurement for work and energy. If you take say a bag of sugar that has a mass of one kilogram and lift it a height of 1 metre you will have done 9.8 joules of work.. You will have given it an extra 9.8 joules of gravitational potential energy. (The untidy number 9.8 arises because the force exerted on the bag of sugar by the earth's gravity is not a whole number.)

The joule is the unit of measurement of work and energy. It is named after an English scientist Dr Joule of Manchester, who, in the nineteenth century, investigated the relationship between work and heat.

Power

The rate at which we do work is called "power". Power is measured in units called watts .The symbol for the unit is W. If you do 1 joule of work in one second you have applied a power of 1 watt. The watt is named for James Watt who in 1765 made changes to the steam engine that were the basis for all future development.

One watt is not very much power. It is about the power needed to light one torch bulb.

The engine of the Mitsubishi Magna has a maximum power output of 120,000 watts, that is, 120 kilowatts. In other words, it can deliver a maximum of 120,000 joules of energy in one second. This car has a mass of about 1600 kilograms

If the engine of a Magna were to deliver constantly 120 kilowatts, of power, then, in the absence of friction, the engine would accelerate the car to 100 kilometres/hour in about 5 seconds.

We know the Magna cannot accelerate that quickly. Something else must be happening to the energy

In fact, the engine cannot deliver 120kW constantly during acceleration from a standing start. And, as we shall see, some of the energy is wasted.

CONSERVATION OF ENERGY

Energy cannot be either created or destroyed. It can only be converted from one form to another.

When we do mechanical work on something, say a trolley, by pushing on it and moving it, we will have changed its speed or its height above the earth.. We will have therefore changed its mechanical energy. The energy we used has been given to the trolley. But, the total amount of energy is unchanged.

Energy Conversion - The principle Illustrated.

Consider the generation of electricity using a steam turbine. Fuel in the form of, say, oil or coal, or natural gas is burned in a furnace. The air used for combustion, together with the gaseous combustion products, become very hot. The furnace is open to the atmosphere so the gas pressure remains close to atmospheric pressure. Chemical energy has been converted into heat.

To do mechanical work we need gas at high pressure. So, these “flue gases” are passed around the outside of a complex of tubes carrying the working fluid, most usually water. The complex is known as a boiler. The water passing through the boiler is heated by the gases and evaporated to form steam. The water is made so hot that the steam above it is at high pressure. The heat energy from the flue gases has been transferred into the water entering the boiler and converted to heat and pressure (that is, enthalpy) in the steam.

The steam itself is then drawn out of the boiler and further heated by the flue gases in a “superheater”. More heat energy is converted into enthalpy.

The steam is then released to pass into the nozzles of the turbine. The velocity of the steam increases as it passes through the nozzles. Some of the steam’s enthalpy is converted in kinetic energy of the steam.

The pressure and kinetic energy of the steam coming out of the nozzles then force the turbine wheel to turn. They exert a torque. However, the kinetic energy of the turbine does not change. Once it is running its speed remains constant. The energy of the steam is not converted to the energy of the turbine.

The turbine is connected to an electric generator - a magnetic rotor which spins inside a coiled conductor so as to generate electric current in any attached circuits. This requires energy continuously. The generator takes energy directly from the steam by means of the turbine.. The temperature and pressure and velocity of the steam decrease as they pass through the turbine.. Much of the enthalpy and nearly all the kinetic energy of the steam are converted into electrical energy.

If you were to read further on this subject you might discover what happens to the heat left in the flue gases and the enthalpy left in the steam. Try to find out what a condenser does.

Energy Conversion - The Inevitability of Losses

In any real machine, some energy always leaks out and passes from the machine to the outside world. This leakage is called "loss". This principle is most easily seen in machines which use heat to do work. That is, for machines, such as the automobile engine, which convert heat into mechanical energy.

Most frequently the lost energy becomes heat or kinetic energy in the atmosphere. It may eventually warm the earth or leak into outer space. The lost energy is not concentrated (the increase of temperature of the atmosphere is very small) and for that reason it is very difficult to recover the lost energy to use it again quickly. It is low grade energy. Although the energy has really been converted not destroyed, it appears to have been consumed.

Efficiency or What You Get - vs -, What You Pay For It.

Engineers are very concerned with how much useful output they can get from a machine compared with the input to the machine. In calculating efficiency it is best if input and output are both measured in the same terms. For example, how much useful energy comes out for the total energy that goes in. Efficiency is defined as useful output divided by input and is usually expressed as a percentage. For example, if you have switched on a 75 watt light bulb in your house, the electric power station will do 75 joules of work on the bulb every second. The bulb will give out 1.5 joules of light every second. The bulb's efficiency as an illuminator is 1.5watts/75watts = 0.02 or 2%.

What would be the bulb's efficiency as a heater?

Energy Consumption in the Automobile

The automobile provides us with a good illustration of losses and efficiency.

Some energy is wasted inside the automobile engine itself. The thermodynamic efficiency of combustion and of the pressure-volume cycle in the engine is only about 30% or 40%. The lost energy passes out through the exhaust pipe to atmosphere; or through the cylinder walls to the cooling water and thence to atmosphere via the radiator.

At the same time, energy is being used to overcome the internal friction in the engine. This energy is also passed to the atmosphere through the cylinder walls and also through the engine oil and the sump and oil cooler.

Next, some energy is taken for what are called "ancillaries". These include battery charging, air-conditioning, power steering, lights, and even the radio.

The remaining energy is available to do mechanical work propelling the automobile. This work is done against friction. Friction arises from:

rubbing surfaces in the gearbox, clutch and differential

rolling resistance of tyres

aerodynamic drag

This energy is converted to heat or kinetic energy in the atmosphere. It becomes low grade energy. It is lost.

The energy used in overcoming friction depends on the speed of the engine and of the automobile.

If the automobile is travelling on a level road at constant speed these processes will have converted all the energy available from the fuel used.

If the driver now presses on the accelerator pedal the amount of fuel and air entering the engine will increase and there will be energy left over after the conversions through friction and ancillaries. This will be available for changing the kinetic and potential energy of the car (including the rotational kinetic energy of the engine, transmission and wheels.) This is called "excess energy". The excess energy available in unit time could be anything from zero to 25% of the chemical energy of the fuel used in that time.

The excess energy available in unit time is commonly known as the "excess power". The speed of the car will increase until all the excess power is used by the increased friction at higher speed, or by climbing a hill, or both.

SOURCES OF ENERGY

Basic Sources of Energy

There are four sources of energy available for us on earth to use:

the chemical potential energy of the various substances on earth,

the gravitational potential energy of moveable items especially water, and

the heat from the sun and the centre of the earth.

nuclear energy in the substances on earth

Chemical energy arises from the potential for substances to react with one another and, in so doing, to produce heat or electricity. The chemical potential energy is chiefly available from things which we burn, that is react with oxygen) so that we can use the heat. These are called fuels. Chemical energy has been accumulated in them over a period of time. Some of these fuels we can produce ourselves so they are called "renewable". The renewable fuels include, for example, wood from trees, other dead vegetation, alcohol from the fermentation of grain and gas from decaying garbage.

There are non renewable fuels such as coal, petroleum, and natural gas. These so called "fossil fuels" took so many years to produce that we could not hope to renew them. They are most valuable because their chemical energy is very concentrated as a result of the millions of years they took to make..

Gravitational potential energy is chiefly available to us as water stored on high ground - which we can use directly, as in a water wheel, or indirectly by the production of electricity which in turn is used for our final purpose. Gravitational potential energy is also present in water at high tide and in waves.

Heat is available directly from the sun and also from the molten magma in the centre of the earth.. We call these solar energy and geothermal energy. Solar energy is also present in the form of wind and in the heat contained by the oceans.

Transmission of Energy

Before we can use the energy from these sources, we must move the energy to the place at which we want to use it. This is known as energy transmission, or more usually as power transmission. There are relatively few methods of energy transmission and, before the energy can be transmitted from our sources it must usually be converted into another form, one that is suited to the means of transmission.

It is also usually necessary to have a machine of some sort to connect or "couple" the energy source to the means of transmission. Energy conversion often happen inside such machines. The automobile engine couples the chemical energy of the fuel to the shafts that transmit energy to the wheels. The pedals couple the chemical energy of the rider's leg muscles to the chain and back wheel of the bicycle which in turn transmit and convert this energy to the kinetic energy and potential energy of the rider and bicycle as a whole.

Many ingenious methods have been proposed for coupling the energy of ocean tides and waves and the wind to electrical generators.

The means of transmitting energy include:

rotating shafts

pushing and pulling of rods levers chains belts and ropes

conduction and convection of heat

conduction of electricity

pressure of fluids in pipes.

Students may find it curious to learn that the lifts in many buildings in London and the cranes on the wharves were once operated by water pressure. There was even a railway that was powered by air pressure.

Heat and electricity may also be transmitted by radiation.

Radiation

If you stand in front of a fire or a glowing electric heater, you will feel yourself being warmed. Energy is passing from the fire to you, even though the fire is not touching you. The same thing happens in a bread toaster. The bread is heated by the glowing electric element in the toaster even though they are not touching. When energy is transmitted across open space we call the process "radiation". The fire radiates, the toaster element radiates, the electric heater radiates.

The sun radiates heat and light to the earth. We receive that radiation and make great use of it.

Of course, receiving too much of the sun's radiation can harm us. We may be sunburned or our eyes may be damaged. The sun's radiation can have detrimental as well as beneficial biological effects.

Sometimes the radiators are called transmitters. Television, radar and radio transmitters, including mobile telephones, radiate and pass information by the way in which they transmit energy across open space. X ray machines do the same. Receiving too much of this sort of radiation can also have a detrimental biological effect.

Sometimes it is useful to think of the energy being transmitted as a beam or ray. Sometimes it is useful to think of the energy as being contained in what scientists called an “electromagnetic field” spreading out from the radiating object.

NUCLEAR ENERGY

Atoms

Most of our scientific theories about the world fit together very well if we believe that everything in the world is made up of small bits called atoms.

An atom is the smallest bit of matter that has distinctive characteristics. There is a different atom for everything. So we speak of oxygen atoms and atoms of iron. Most everyday things are made up of atoms of more than one sort. Rust, for example, is a mixture of atoms of oxygen and iron and hydrogen.

We also say that an atom itself is made up of smaller particles called "sub atomic" particles. These have names like neutron, proton and electron. The protons neutrons and electrons in an atom are the same as those in any other atom. Protons and neutrons are in a part of the atom called the nucleus.

Radioactivity

Under certain circumstances atoms do break up or "decay". When they do, they emit sub atomic particles and they radiate some energy.

Quite a few substances have atoms that are a little unstable - the atoms are prone to breaking up spontaneously.

When first noticed it seemed as though the substances were naturally radiating energy. They were called "radioactive".

Some substances were found to emit "radiation" when light was shone on them - that is when they absorbed radiant energy. Some re emitted light. These were called fluorescent and phosphorescent. M. Becquerel investigated this effect in uranium in 1896.

Mme and M Curie investigated the matter more deeply and concluded that these effects were "atomic". We have since discovered that there are many radioactive substances. Radium, plutonium and uranium are only three.

Substances were found which conducted or even generated electricity when a light was shone on them. These effects are called "photoconductive" and "photovoltaic" or simply "photo electric".

Atoms of most substances will break up if they are hit by a speeding proton, electron or neutron. Lord Rutherford demonstrated this in 1919. He broke up atoms of nitrogen and they turned into oxygen .

When the early scientists observed radioactive decay, they called the emissions alpha, beta and gamma rays because they did not know what they were.

We now know the particles mainly emitted from a decaying atom are neutrons, neutrons and protons joined together called "alpha particles", and electrons also known as "beta particles".

The radiant energy which is emitted at the same time may be heat or light or X rays and similar radiation called gamma radiation.

Even though the emission of particles from a decaying atom is not exactly the same as radiation of say heat and light, we continue to refer to the emitted particles and energy as "atomic radiation", or "nuclear radiation” if the particles come from the nucleus of the atom.

Effects of Atomic Radiation

The speed, size and mass and the electrical characteristics of these particles, and the intensity of the gamma rays, determine what happens when they hit something.

Alpha particles have a mass of only about 6.6 x 10-27 kg, but they are still 7400 times heavier than beta particles and four times heavier than neutrons. They are emitted from the nucleus at an enormous speed - about 15000 kilometres/sec .But, they are easily stopped. Because of their size and electrical properties they will only travel about 40mm through air.

Beta particles sometimes have speeds even higher than the alpha particles - as much as 95% of the speed of light for example.

The energy of the emissions is absorbed by any object which stops them. This transfer of energy to an object invariably heats or damages the object - sometimes minutely sometimes grossly..

To estimate the likely effect of absorbing atomic radiation, we measure what is called the radiation “dose” or “dosage”. An instrument for doing this is called a “dosimeter”. Often it takes the form of a badge made of photographic film. You may see such badges being worn by laboratory personnel.

The overall effect of the nuclear radiation will depend upon how big an object the radiation hits.

So we calculate the dose as the amount of radiation received for every kilogram mass of the object being irradiated. We say an object has received a dose of 100 rads when one joule of energy is absorbed by one kilogram of mass.

Each type of atomic radiation causes its own degree of damage to living tissue. The effect also varies from one type of tissue to another. So the number of rads absorbed does not tell the full story for biological effect. Ten rads of neutrons does not have the same effect as ten rads of gamma rays or ten rads of alpha particles.

If we take into account the relative biological effect of the type of radiation, we can calculate a “rad equivalent for man” called a “rem” and use this to measure doses of radiation and to specify the dose of radiation that is safe. Suppose one piece of tissue absorbed 10 rads of slow neutrons and ten rads of alpha particles. The effective dose would be about 50 rem for the neutrons and 120 rem for the alpha particles that is a total dose of 170 rem

Nuclear Fission

When the nucleus of very large atoms, such as those of uranium are hit by speeding neutrons they do not just emit alpha particles and neutrons. The atom actually splits into two pieces. It is said to undergo “fission”

The two pieces are each much larger than an alpha particle.. They are in fact atoms of other substances. They travel at enormous speeds after fission. When a uranium atom undergoes fission it splits into atoms like barium and krypton. These are called “fission products”.

Now one atom of barium together with one atom of krypton is lighter than one atom of uranium. It seems that there is some mass missing and indeed there is. The mass has been converted into energy. This is where

Einstein's famous equation $e=mc^2$ applies. It is used to calculate the amount of energy that corresponds to the missing mass. The mass is converted mainly into kinetic energy of the fission products.

The atoms into which uranium has been split are themselves radioactive. They decay and give off about 7% more energy.

The kinetic energy of the fission products is transferred to what ever objects they run into and, generally speaking, will heat those objects.

The Chain Reaction

When a uranium atom is hit by a speeding neutron and undergoes fission it also emits several neutrons. These neutrons are then able to cause fission in adjacent atoms. Fission in the first atom will lead to fission in two or three adjacent atoms, and each of these will cause fission in two or three more atoms. So the whole process is a self-sustaining chain reaction. Unless it is controlled it will go faster and faster. More and more mass will be turned into energy faster and faster. More and more heat will be generated and eventually there will be an explosion. This is the principle behind the atomic bomb.

The Nuclear Reactor

There are several types of uranium atoms. Naturally occurring uranium contains about 7% of the less stable type U_{235}, with the remainder being mainly U_{238}. U_{235} can be split by slow moving neutrons. The most abundant type, U_{238} absorbs slow and medium speed neutrons and turns into $Plutonium_{239}$ which is also fissionable material. The fission chain reaction proceeds more quickly if all the neutrons emitted cause further fission. This means some of the neutrons must be slowed down so as to be sure they find and collide with the small amount of U_{235} that is present. Alternatively the natural uranium must be "enriched" by the addition of more U_{235}..

A nuclear reactor is a means of allowing the chain reaction to proceed in a steady controlled way with balanced production and absorption of fast and slow neutrons. Curious students may like to use the library to find out what is meant by the terms "fast reactor' and "thermal reactor".

. The fissionable material in a reactor is contained in "fuel elements". The substance which slows down the neutrons is called a "moderator". Graphite, beryllium, boron, and heavy water have been used as moderators. A fluid is used to conduct away the heat energy to a place where it can be dissipated or used.

When the fuel elements are no longer able to generate energy from fission in the reactor they are said to be "spent". Spent fuel rods are still radioactive and present a health hazard unless disposed of carefully. The disposal of this material is still a matter of contention, even though nuclear reactors have been running for many years.

Nuclear reactors are found in scientific research establishments and nuclear power plants. They are also used for nuclear weapons production. Reactors used to turn abundant types of uranium into rare highly active material such as plutonium are called "breeder reactors".

In a nuclear power plant, ie a nuclear fuelled electricity generating station, fission heats the fuel elements, the fuel elements heat water, the water turns to steam, the steam drives turbines and the turbines drive alternators to produce electricity.

Relative Efficiency

Nuclear reactors with steam boilers connected to steam turbines connected to alternators are about 30% efficient. If the steam is generated by burning coal, the efficiency is about 40%. This difference arises because of difficulties experienced when operating reactors at the high temperatures needed for best efficiency.

Suppose there are 500,000 dwellings in Brisbane. If they all had a 1.2 kilowatt electric kettle switched on at breakfast time they would be drawing a total of 600megawatts from the generating station - just for the kettles. If the kettles were left on all day, the power station would use 6okg of uranium based fuel. To obtain the same effect with a coal fired power plant would require about 3,300 tonnes or 3.3 million kilograms of coal per day.

Of course, coal is very much more abundant than uranium. The cost of the coal would only be about 2.5 times the cost of the nuclear fuel. But, consider. All that coal must be transported from the mine to the power station, unloaded, stored, and then loaded into the furnaces. Then all the ash must be carried away. The nuclear powered generator would only need fuel supplied say once per year. The coal fired station would require a continuous stream of trains, as well as bulk stockpiles at the mine and the generating station.

Does this suggest that, before beginning to discuss energy economics, one should define the boundaries of the problem so that one compares like with like? Perhaps one should draw a diagram of the whole system and identify the costs (ie price and losses due to inefficiency) at key points in the diagram.

ELECTRICITY

ELECTRICAL RESISTANCE

Ohm's Law

The 19th century experimenter Georg Simon Ohm discovered that in an electrical circuit the current flowing through a conductor was directly proportional to the potential difference between the ends of the conductor.[1] That is, potential difference = constant x current. This relationship became known as Ohm's law.

Using modern symbols[2] we would write this as:

$$V = \text{constant } xI \quad \text{or,} \quad E = \text{constant } xI$$

The constant of proportionality is known as "Resistance" and it is customary to give resistance the symbol R..

Potential difference, and its special form "electromotive force" are described in units of volts with unit-symbol V. For example E = 12V

Current is described in units of amperes, colloquially "amps", with unit-symbol A. For example I= 3A

Resistance is described in units of ohms, with unit-symbol Ω. For example R=998Ω

One ohm is defined as the resistance of a conductor in which a potential difference of one volt causes a current of one ampere to flow.

It is common to see decimal multiples of ohms, and these are always written using the ternary powers of ten such as kilo and mega. For example R=15kΩ or R=2.7MΩ.

Correspondingly, it is common to see decimal submultiples of volts and amps, such as millivolts mV, and microamps μA.

The correct use of upper and lower case letters in the unit symbols is most important. Upper case is used when the unit-symbol is derived from a person's name. Lower case is used elsewhere.

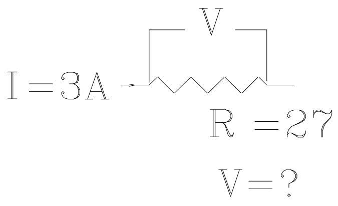
---

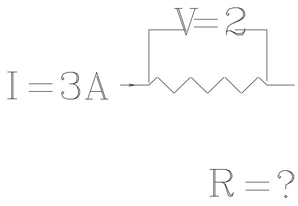
1 Teachers and students may like to reflect on how Ohm measured current and voltage. Had the d'Arsonval Galvanometer, which forms the basis of the voltmeter and the ammeter, been invented and does its principle depend on the validity of ohms law?

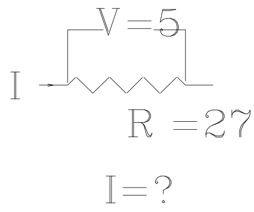
2 It is more usual to use E as a symbol for potential difference because V is used as the abbreviation for the unit of measurement. To write V=12V would be very confusing.

The most common form in which Ohm's Law is expressed is E=IR

This can be transformed to $I = \frac{E}{R}$ or $R = \frac{E}{I}$







Resistivity

The electrical resistance of components in electrical circuits depends upon their length, their cross sectional area, and upon a property of the material from which they are made known as the “resistivity” of the material.

Materials commonly used as conductors typically have resistivity of about 10^{-7} ohm.metres.

Accordingly, the resistance of conductors is usually regarded as negligible unless they are comparatively long.

Materials used as insulators typically have resistivity of about 10^{13} ohm.metres.

There is a class of materials known as semiconductors which have resistivity between 1 and 2500.

The resistivity of most materials, with the notable exception of carbon, increases as temperature increases and decreases as temperature decreases. There is thus the prospect that zero resistance might be achieved at very low temperatures. This in fact is achievable with lead when cooled to below -265C. These materials are known as superconductors and the phenomenon is called superconductivity. In 1988, new materials were discovered which were superconductive at -150C

Students might like to consider what applications there are for materials that have no electrical resistance - remembering that they are expensive and must be kept very cold. Would they be useful, for example, in a train supported by magnets rather than wheels - “magnetically levitated”.

When current passes through a resistance, power is dissipated. The resistor becomes hot and passes heat to surrounding components and to the atmosphere. As the resistor becomes hot its resistivity and therefore resistance increases and the current it draws reduces.

Most semiconductors are very sensitive to temperature. Their uses thus include making temperature dependent resistors called “thermistors”

Semiconductor material can be chemically formulated to have a high resistivity in one direction and a low resistivity in the other, and resistivity that changes with applied voltage. This is the foundation of semiconductor electronics. The word “transistor” is a contraction of “transfer resistor”.

The resistivity of some semiconductors changes when they are exposed to light. These materials are used to make “light (or photo) dependent resistors” and “photo transistors”.[3]

Resistors as Circuit Elements

In some electrical circuits resistance is present incidentally to the operation of the circuit - usually as a natural consequence of having to use relatively long lengths of conductors. Power transmission lines are an example. So too is the resistance of the coiled conductors used to make electromagnets in solenoids, motors, alternators, and moving coil meters.

In some circuits, conductors are chosen purposefully for their high resistance. This is so in heating elements. In electronic circuits, conductors with high resistance, components known as “resistors”, are used

---

[3] And also photo sensitive diodes.

to adjust voltage and current at various points in the circuit. In later work, you will discover that they are used for purposes such as limiting the current flow through light emitting diodes, setting bias conditions for transistors, adjusting the charging rate of capacitors, and adjusting feedback in amplifiers.

In choosing a resistor for these latter purposes, one must consider not only the value of the resistance needed to do the job, but also the accuracy needed, and thus the tolerable inaccuracy in the resistor.

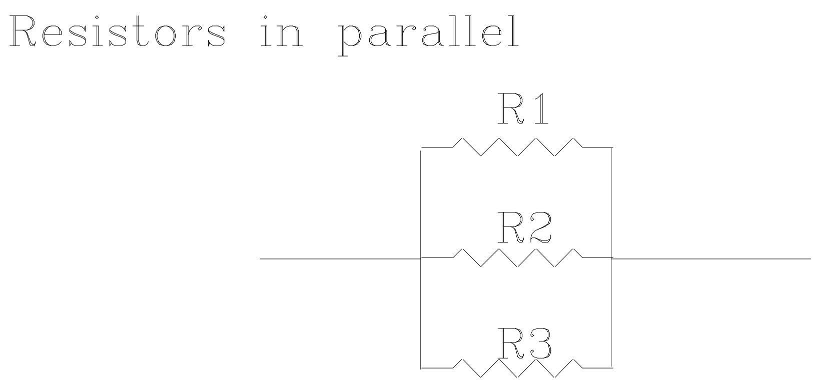
It is also necessary to estimate the power which the resistor will be asked to dissipate and to choose a resistor of a size and material that will dissipate that power safely.

Resistors used in electronics are marked to indicate their resistance, their tolerance grade (typically 1% or 5%) and their maximum power dissipation.

Sometimes it is useful to be able to calculate the equivalent total resistance of a combination of resistors.

When resistors are connected so that the same current is flowing through all of them, the resistors are said to be "in series".

When resistors are connected so that the current flowing is divided between them the resistors are said to be "in parallel".

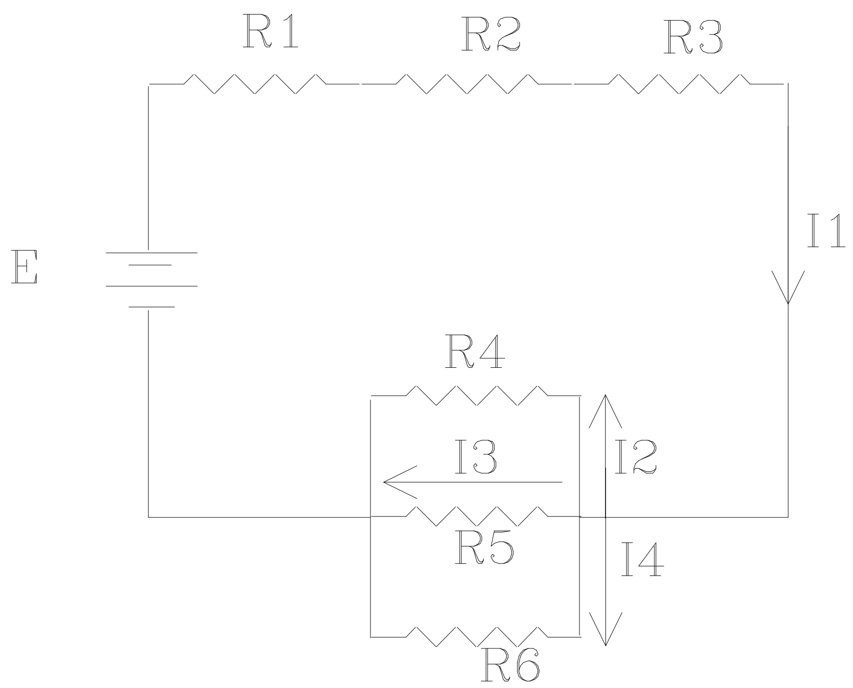


When resistors are combined in series, as shown, the total resistance of the combination is given by: $R = R1 + R2 + R3 + \ldots$

When resistors are connected in parallel, as shown, the total resistance of the combination is given by: $\frac{1}{R} = \frac{1}{R1} + \frac{1}{R2} + \frac{1}{R3} + \ldots$

Note that when resistors are combined in parallel the voltage drop across each resistor is the same.

It is common sense that in a circuit the algebraic sum of the currents flowing into a junction is zero. A second useful rule is that in any circuit loop, the algebraic sum of the emfs and potential drops is zero.[4]



In the diagram above, Ohm's law and the rules for adding resistance combinations tell us that:

$$E = I1(R1 + R2 + R3) + I1 \times \frac{R4R5R6}{R4 + R5 + R6}$$

The rules for circuit junctions and loops tell us that:

Summing the currents at the junction:

$$I1 - I2 - I3 - I4 = 0 \text{ and}$$

Summing the voltage drops around the circuit:

$$E - I1 \times R1 - I1 \times R2 - I1 \times R3 = I2 \times R4 = I3 \times R5 = I4 \times R6$$

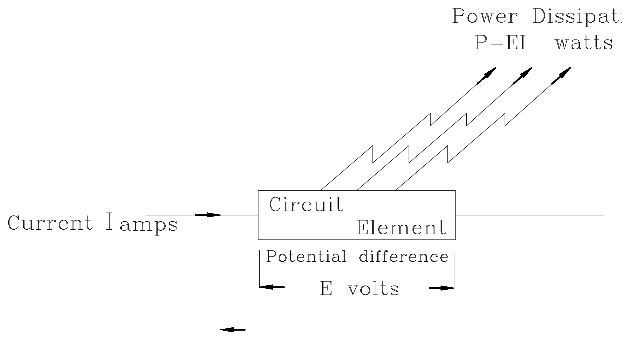
These equations may be used to find the voltage and current at any point in the circuit

Your teacher may like to go on and describe the use of resistances in a voltage divider, the concept of measuring resistance using a potentiometer, and the derivation of the formula for resistances in parallel.

4 These are known as Kirchoff's !st and 2^{nd} laws.

Electrical Power

By experiment, we find that if a current travels along a conductor between points of different potential, energy is dissipated by the conductor. If a current of one ampere is flowing, and energy is being dissipated at the rate of one joule per second, we say that the potential difference between the two points in the conductor is one volt.



We can turn this idea around. Suppose we look at two points in a circuit and instantaneously measure the current flowing and the potential difference between the points. If the current is one ampere and the potential difference is one volt, we know that one joule of energy is being dissipated per second. Now, the rate at which energy is used is called "power" and the use of one joule per second is a power of one watt. So we can write:

$P=EI$
where
P = the power in watts being used in this element of the circuit
E = the instantaneous potential difference cross the element measured in volts
I = the instantaneous current flowing through the element measured in amperes

Using Ohm's law we can write:

$P=I^2R$
where R is the resistance of the particular element of the circuit.

Remember

Twinkle twinkle little star,
Power equals I squared R

GENERATION AND RETICULATION OF ELECTRICAL POWER

Electricity and Magnetism - Electromagnetism

During the 1830s, scientists began to investigate the relationship between electricity and magnetism which had been discovered by Hans Oersted of Denmark in 1819. An Englishman Michael Faraday and an American Joseph Henry discovered that if a magnet was close to a coil of wire, then when either the coil or the magnet were moved an "electrical pressure" was created between the ends of the wire.

When ever there is an electrical pressure between two places we call this pressure a "potential difference". One place, one end of the wire for example, is at a higher electrical "potential" than the other end.

Scientists already knew that if a place of high electrical potential was connected to a place of low electrical potential, by say a piece of wire, then an electric current would flow from the high potential point to the low potential point.

Sure enough that was exactly what happened when a wire was connected to the ends of the coil near the moving magnet. The potential difference forced a current through the wire.

A potential difference created for that purpose is called an "electromotive force" or "emf".

The volt is the unit of measurement for potential difference and electromotive force, so that pd and emf are very often colloquially called "voltage". We ought not to use colloquial terms in science. The unit is named in honour of Alessandro Volta, probably the first person to make an electric battery.

The unit of measurement for electric current, the amp, is named in honour of Andre Marie Ampere - a French scientist who devised theories of electromagnetism. Current is some times referred to as "amperage".

It is useful to think of electric current as a flow of electrically "charged particles" These particles flow from a place of higher, say positive, potential to a place of lower, say negative potential. The greater the potential difference the more the current. The more the current, the more the charge that flows in a given time. [5]

Earth Potential

The earth is so big that to all intents and purposes its potential cannot be raised by electricity flowing into it as a result of our activities. The earth is regarded as a place of zero potential for man-made electrical devices.

In fact the earth and its atmosphere are a large electric generator powered by the sun.

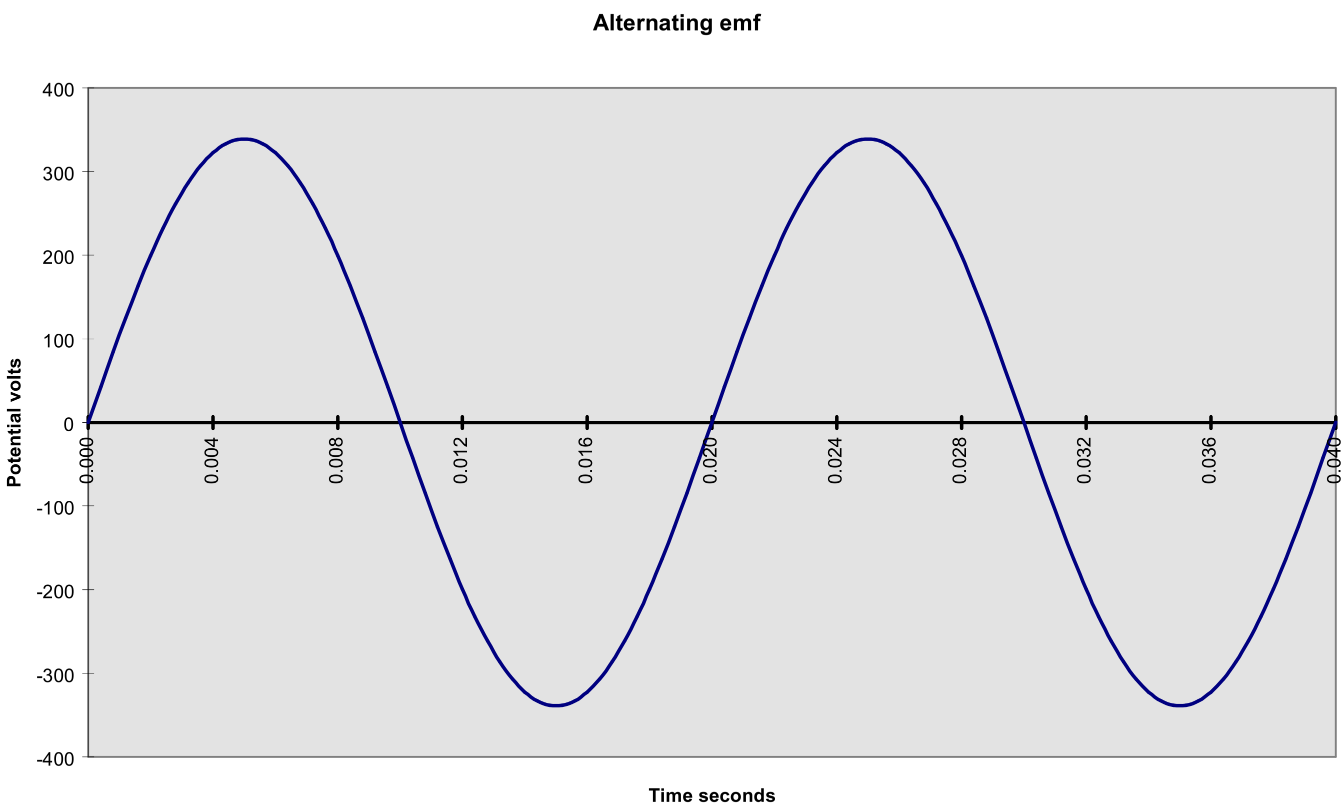
Alternating Current

Let us return to the magnet and coil moving relative to one another. To keep the magnet and the coil close together but still moving it was necessary that the motion be rotation. Scientists then found that as the magnet (or the coil) turned through each full revolution, the potential at the ends of the coil wire swapped over. Each end was alternately high and then low.

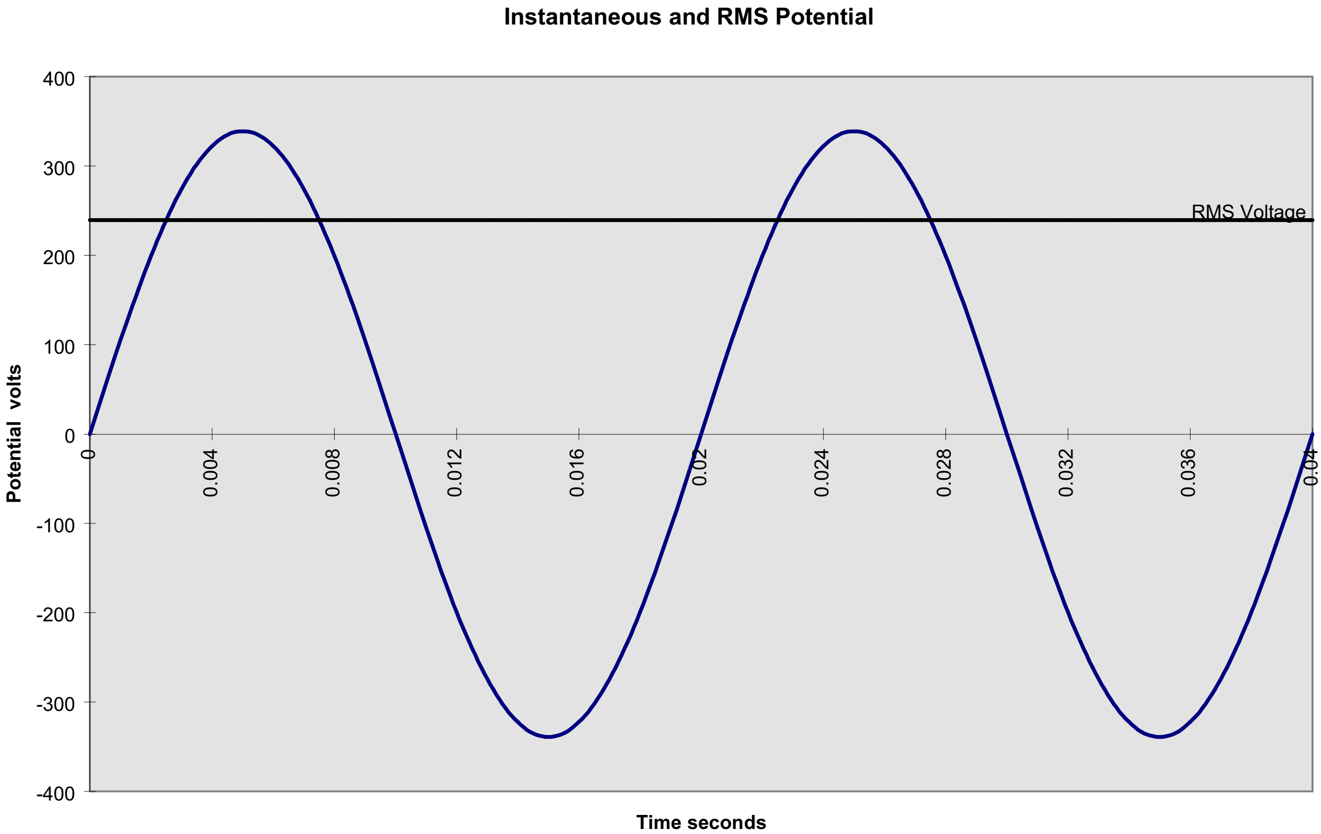
In consequence, the current in the circuit formed by the external loop of wire and the coils themselves, alternately flowed in one direction then the other. This is known as "alternating current" or AC. The changeover is not sudden but is a smooth change from zero up to maximum in one direction, down through zero again and then to maximum in the opposite direction and back to zero. The machine with magnets and coils which produces alternating emf and current is known as an "alternator".

---

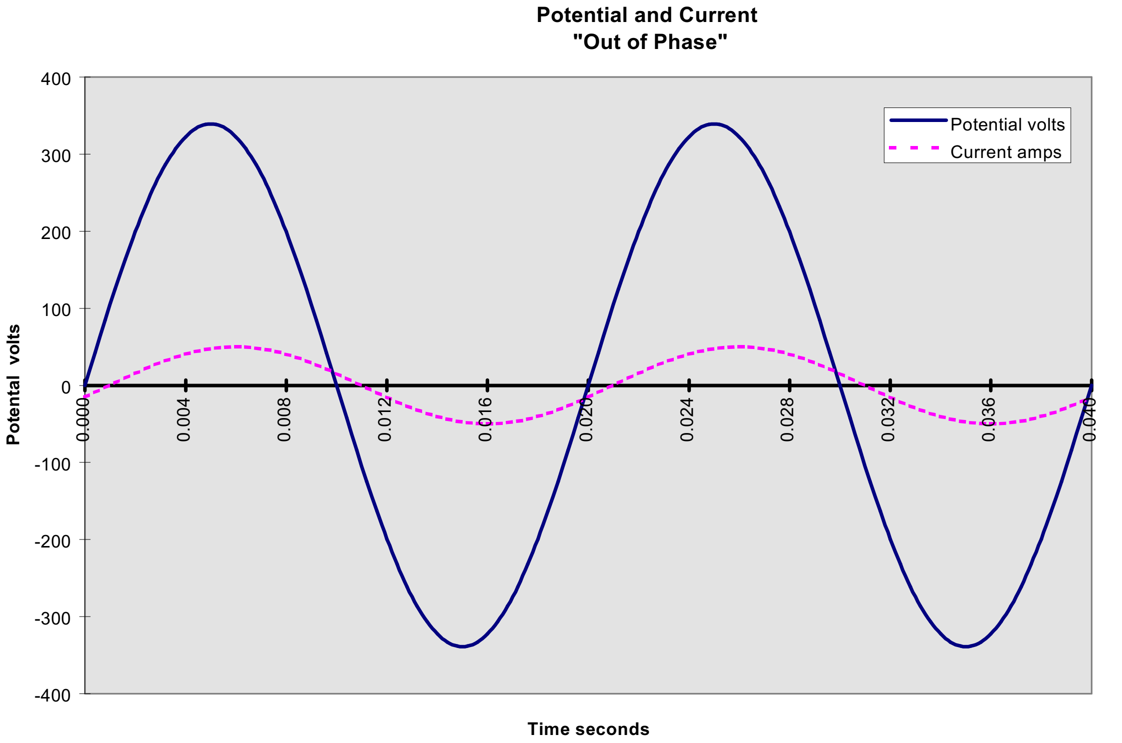
[5] Note well; it is customary always to describe currents as though they consisted entirely of the movement of positive charges, irrespective of the subatomic actuality.

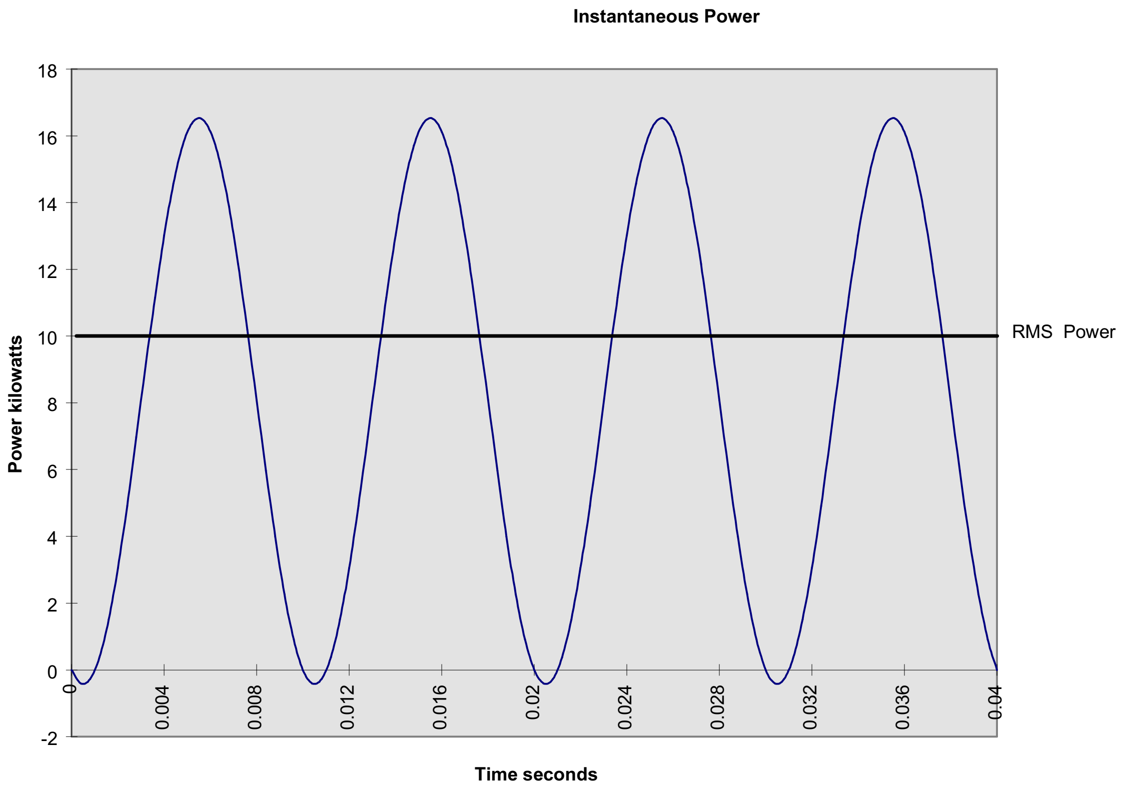

Alternating emf
Potential volts
400
300
200
100
0
-100
-200
-300
-400
0.000
0.004
0.008
0.012
0.016
0.020
0.024
0.028
0.032
0.036
0.040
Time seconds

Because the current and voltage are continuously changing, we usually describe alternating current in terms of the frequency of alternation, the average current and the average voltage. We use a special type of average called the root mean square or RMS. For example our electrical outlets in houses in Australia supply about 240V RMS at 50cycles per second, also called 50 hertz.



The formula for power P=EI uses instantaneous voltage and current. In AC, the RMS voltage and the RMS current, which we use to describe AC, are not the instantaneous values. We cannot use them in the formula. Instead, with RMS values, we must use the formula $P=EI\cos\phi$.. $\cos\phi$ is called the power factor and in usually between 0.85 and 1.0. It allows for current that is "out of phase" with the potential. That is where the peak current occurs after the peak potential. This happens in many circuits carrying AC





Direct Current

We know that it is possible to obtain electrical energy from chemical energy using the device we call a battery. When we do that, the current in any external circuit we may attach to the battery does not alternately flow in one direction then the other. It flows always in the one direction, out from the positive terminal of the battery and back in at the negative terminal This is called “direct current” or DC.

It is possible to generate DC electromechanically much in the same way as AC is generated. All that is needed is a switching device which swaps the ends of the alternator coils over just as the direction of the current is about to change. (This switch is called a commutator). DC can also be made from AC using a device called a rectifier. The current produced in these ways is not exactly constant DC but can be made close enough to DC for practical purposes.

Advanced Note on RMS Power

Some students may notice that in the diagram showing instantaneous power and RMS power there are small sectors of “negative power”. The explanation is that the circuit element under consideration is returning energy to the rest of the circuit, as distinct from absorbing or dissipating it.

This leads on to a more detailed treatment of power. The **average power** being dissipated over a cycle, by a resistor or transducer say, or perhaps being absorbed by an inductor or a capacitor, is

$$P = \frac{1}{T}\int_0^T EIdt$$ where T is the period of the cycle.

If both voltage and current are sinusoidal with frequency *f*, $E = E_m \sin 2\pi ft$. Thence, writing $2\pi ft = \theta$

$$P = \frac{1}{2\pi}\int_0^{2\pi} E_m \sin\theta \bullet I_m \sin(\theta - \phi)d\theta$$ where subscript m indicates peak values

and ϕ is the phase angle by which the current lags behind the voltage.

Recalling that $\sin(\theta - \phi) = \sin\theta\cos\phi - \cos\theta\sin\phi$ And proceeding with the integration , yields

$$P = \frac{E_m I_m}{2}\cos\phi$$

Recalling also that for a sinusoidal variable the RMS value of the variable is $\frac{1}{\sqrt{2}}$ times the peak value and substituting above, we see that the average power being dissipated or absorbed is :

$$P = E_{rms} I_{rms} \cos\phi$$

Generation and Reticulation of Electricity

For reasons which will be discussed later, our industrial and domestic electricity supplies are based on AC.

The alternators which provide our industrial and domestic electricity supply have three sets of coils. They provide what is known as a “three phase” supply. The purpose is primarily to increase the efficiency of the alternator.

The coils are connected like a star. The point where the three coils are connected to one another is called the neutral point. At the power station, the neutral point is connected to the earth and it remains at earth potential.

There is a large potential difference between the neutral point and the other or “active” end of each coil. This is an alternating potential difference.

Every house connected to the power station is served by a wire from the active end of one coil and a wire from the neutral point. This is called a single phase supply.

(Some houses, and most factories are connected to all three phases.)

Inside the house, at the switch-board, the neutral wire is at about earth potential. The potential of the active wire is an average of 240 volts different from the neutral wire.

Neutral and active wires run from the switch-board to every light fitting and power-point in the house. When an appliance is plugged in and switched on it is connected from the active wire to the neutral wire, thus completing a circuit extending from the power station and through which current can flow.

As a safety precaution, switches used to control appliances work by disconnecting either the active wire or both the active and neutral wires from the appliance. However, because of the way they are built and installed, light fittings which have switches some distance away remain connected to the active wire even when switched off.

Efficiency in Electricity Reticulation

The power lines which carry electricity to our factories and houses are made from materials which conduct very well - either aluminium or copper. Nevertheless they do have resistance. When current flows through them they dissipate energy - according to the formula $P=I^2R$.

These power lines also have characteristics called inductance and capacitance that increase the dissipation of energy somewhat when alternating current is being transmitted. Moreover, AC power generation requires more insulation than DC. To supply a given amount of power, the maximum voltage which AC power lines and generators must stand without sparking is about 1.4 times the value for DC. So DC transmission would seem to be most efficient.

Next, if the customers of our electric generating station are going to use energy at the rate of say 10 million watts, we have the option of supplying them power at say 100,000V at 100A or 100,000A at 100V. The higher the current the more the power dissipated by the transmission line, so we choose to reticulate electricity at very high voltages.

Now while it is possible to produce high voltages with DC, this has not always been so. On the other hand, it has been possible to create high voltage AC for over one hundred years. This is done by using a device known as a “transformer”. Transformers can be used to increase or decrease the voltage of AC but not DC.

It is for these reasons that we reticulate electrical power as AC and at very high voltage. But note that some DC reticulation is used in special cases.

The transformer consists of two coils of wire wound on a shared iron core. They are enclosed in oil-filled casings so that they can be more easily cooled. There will be one somewhere near your school.. What is its function?

ELECTRICAL SAFETY

Physiological Effects of Electricity

If you simultaneously touch a wire at high potential and either another wire at low potential or the earth, a current will flow through your body.

The human body is very sensitive to externally supplied electric current because it interferes with nerve processes especially those controlling the heart.

Currents between 0.01 amp and 0.1 amp can cause convulsive muscle action and considerable pain. Higher currents will cause burning. Smaller currents can cause heart trouble.

Nevertheless, some electric currents can be physiologically beneficial. These are very small currents at extremely high frequency of alternation.

The actual current which flows through the body will depend on the potential difference experienced and the resistance of the path between the parts of the body to which it is applied.

Seventy volts applied between the right and left hands will cause a serious shock.

The average potential difference between the neutral and the active wires in our houses is 240volts. The maximum potential can be as much as 360 volts. The frequency at which the potential alternates, 50 times per second, combines to make this a very dangerous supply system. Electric shocks can be fatal.

Is there a poster in your school showing how to safely help someone who has received an electric shock?

What sort of fire extinguisher should you use on a fire caused by electricity?

Safety Earth.

If a fault develops in an appliance and the active wire, or something connected to the active wire, touches the frame of the appliance, the potential of the frame will become the same as the active wire - considerably different from neutral or earth. If the frame is metal, any one touching it would receive a severe electric shock, but for the precautions which safety authorities insist on.

The precaution is that any appliance with external parts that might conduct electricity must have those parts connected to a wire known as a safety earth. The safety earth connects the frame of the appliance, through a three pin plug and socket on the wall, to the earth at the switch-board of the house. The effect is to limit the potential which the frame might reach and to which the user might be exposed..

Fuses and Earth-Leak Detectors

If an appliance is damaged internally it may draw an unusually high current. For example, parts connected to the active wire might touch the neutral wire. There would be an internal “short circuit”. This would also happen if the active wire touched the frame of the appliance, because the frame is connected to the earth zero potential through the safety earth wire. Unusually high currents can cause overheating and fire.

A device to break the circuit in the event of unusually high currents is installed in the switch-board for the house. Most commonly this has been a short length of wire which melts, that is fuses, when required to pass too much current. The device is called simply a “fuse”. The same effect can be achieved using a small switch that opens when asked to pass too much current. These are known as “circuit breakers”.

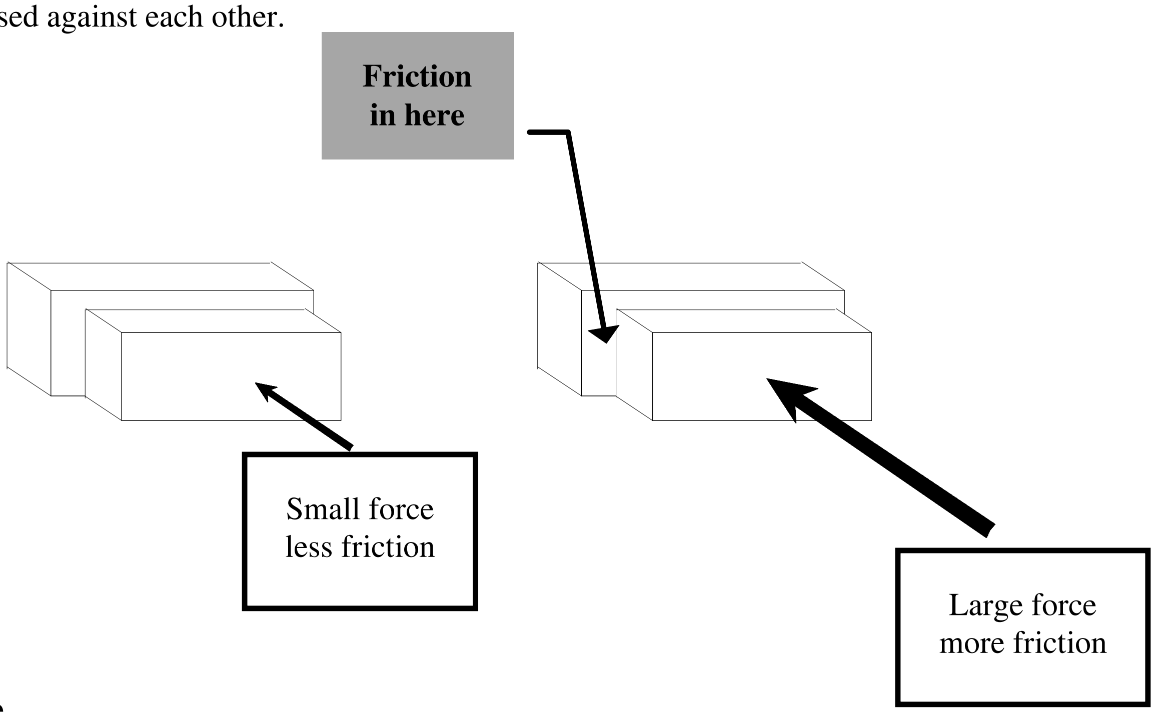
Circuit breakers and fuses have the two effects. They limit the current which flows in the wiring and they improve user-safety by disconnecting the active wire from the main supply if the active wire touches the frame of the appliance. Fuses and circuit breakers do not act quickly enough to completely prevent electric shock to users.

It is possible to install a device known as an “earth leak detector” which disconnects the active wire from the main supply sufficiently quickly to prevent electric shock.

FUSSY FACTS ABOUT FRICTION

RIGID BODY OR SLIDING FRICTION

When two solid bodies that are practically rigid touch one another, there is an interaction between the parts of their surfaces that are in contact. This interaction tends to prevent the bodies from sliding past one another and is known as "friction". It is also called "sliding friction" or "Coulomb[6] friction" after one of the scientists who investigated it. The interaction arises out of the fine structure of the contacting surfaces of the bodies - their molecular nature and their texture. The strength of the interaction depends on how hard the surfaces are pressed against each other.



Frictional Force

In order to do calculations when friction is present we represent the friction interaction by a force acting on each body which we call the "frictional force". We give this force some odd characteristics:

It is only present when something is trying to create relative motion between the two surfaces.

On each body:

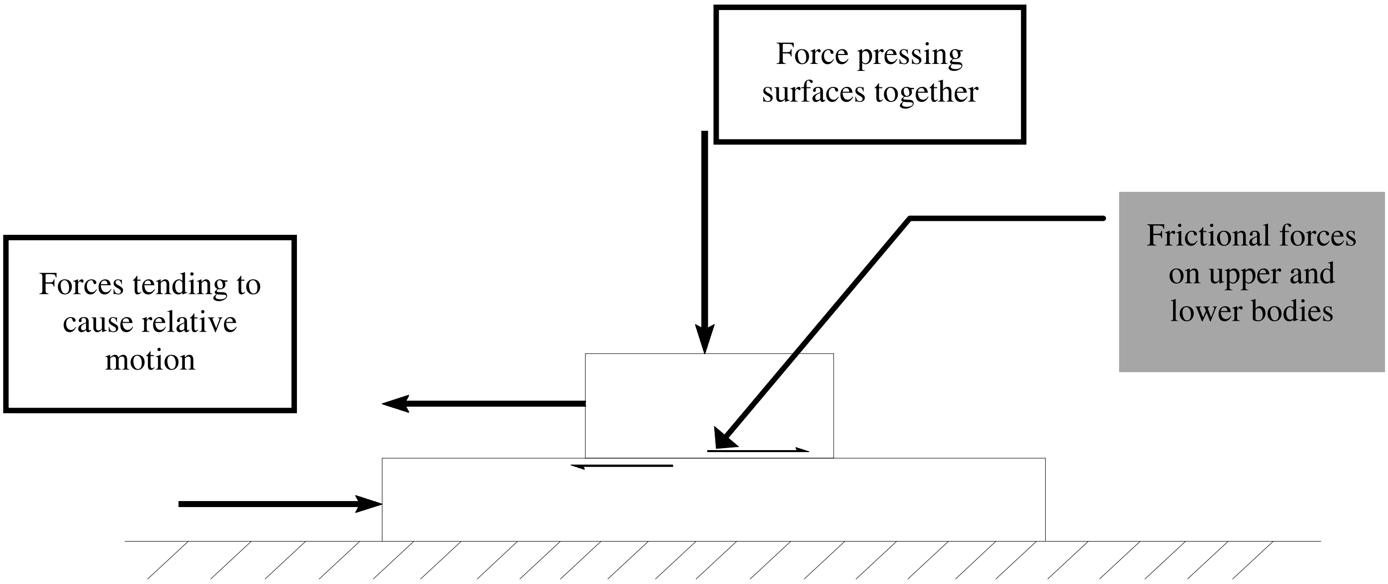
- It acts at the contacting surface where sliding is expected
- It acts parallel to the contact surface and in whatever other direction is needed to resist the expected relative motion between the bodies.
- It is always just large enough to exactly balance whatever force is trying to create relative motion between the bodies - up to a limiting maximum.

---

[6] C.A. de Coulomb 1736 - 1806

If the force trying to create relative motion is greater than the limiting maximum for the frictional force, the bodies slide across one another.

The limiting maximum for the frictional force is called "limiting friction force" or just "limiting friction".



The Laws of Friction

Coulomb discovered that:

- The limiting friction force does not depend on the size of the surfaces in contact
- The limiting friction force depends only on the magnitude of the force pushing the two surfaces together.
- The limiting friction force is directly proportional to the magnitude of the force pushing the two surfaces together.

These rules apply even when the bodies are actually sliding past one another.

Later researchers such as Rennie and Morin, in about 1830, discovered that these laws change at higher contact forces and sliding speeds. Nevertheless we continue to use Coulomb's rules as a good working hypothesis.

The Coefficient of Friction

We have said, the limiting friction force is directly proportional to the magnitude of the force pushing the two surfaces together. The constant of proportionality is called "the coefficient of friction".

We use the formula $F = \mu N$ where:

$F =$ the limiting frictional force
$N =$ the force pressing the surfaces together
$\mu =$ the coefficient of friction

Here are some coefficients of friction you will find listed in various reference books. Circumstances change so much that it is often best to actually measure the friction coefficient for yourself.

Contact Surfaces	Coefficient of Limiting Sliding Friction
bronze on bronze	0.20
bronze on cast iron	0.21
leather on hardwood	0.33
wood on stone	0.40
brass on hardwood	0.48
hardwood on hardwood	0.48
cast iron on hardwood	0.49
cast iron on stone	0.50
leather on cast iron	0.56
brick on stone	0.65

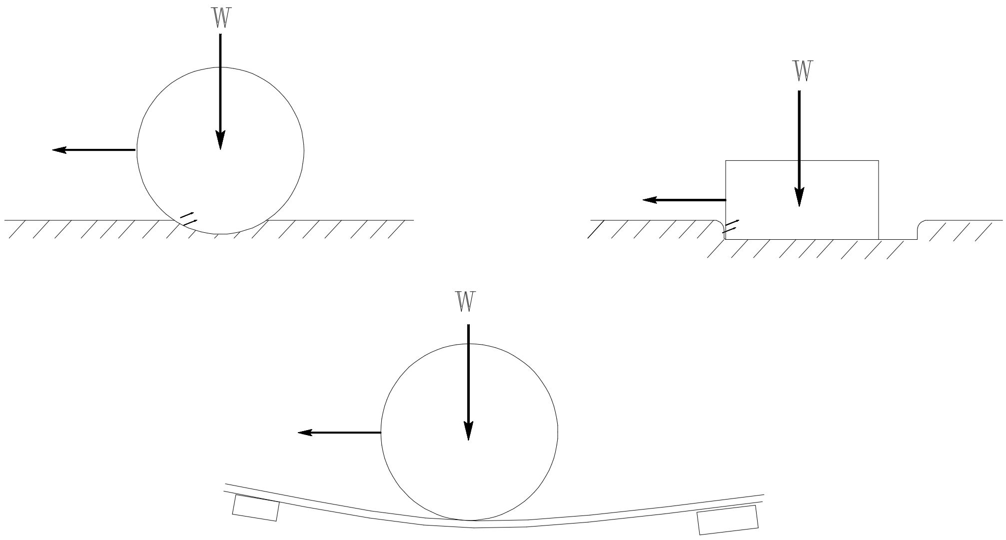
Once the bodies actually start to slide past one another, the coefficient of friction decreases to perhaps one fifth of the value just before the start.

FRICTION BETWEEN NON-RIGID BODIES

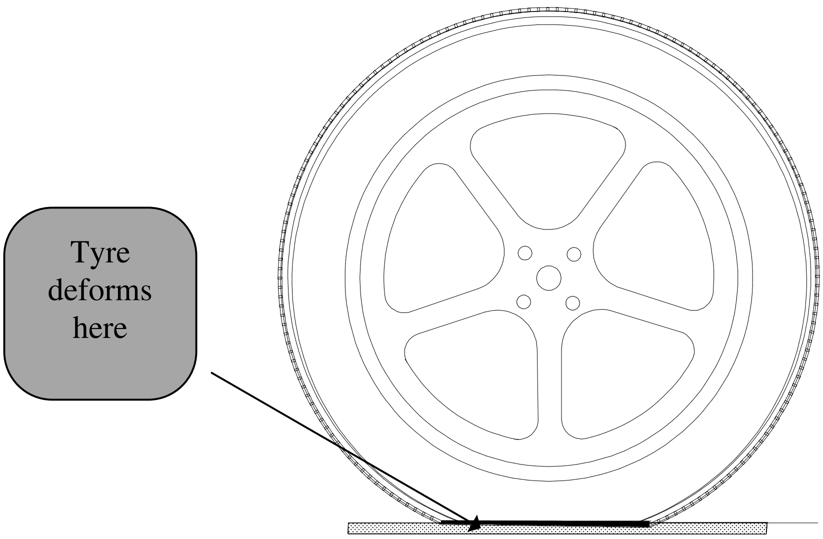
The discussion so far has been about bodies that are practically rigid, that do not deform. That is to say, it has been about situations in which the material of the bodies has been so stiff, or the forces between them have been so small, that the bodies have not noticeably changed shape.

If the bodies do change shape, if one indents the other, Coulomb friction will still be present but there will be a new source of retardation. This will be due to the extra force needed to climb out of the indentation or to flatten out the deformation. This is not friction. But, experimentally, it can be hard to distinguish between friction and the forces needed to overcome deformation. Often the experimental results are wrapped up together.

Examples of Deformation



Another example of deformation



It is better to call the retardation rolling resistance or dragging resistance rather than friction.

FLUID FRICTION

When a gas or a liquid is in contact with the surface of a solid, there is an interaction between the fluid and the surface that tends to prevent them from sliding past one another freely. This is known as "skin friction". The interaction arises out of the fine structure of the solid surfaces and of the molecular nature of the fluid.

This retardation is experienced when, for example, a fluid flows through a pipe, or a thin flat plate slides edgewise through a fluid.

We call the relevant molecular characteristic of the fluid its “viscosity”

We do not know exactly what causes the interaction but there are lots of formulae that have been derived to describe experimental results. We might call these formulae “algebraic models” of fluid friction.

Fluid skin friction must be distinguished from the retardation caused by deformation of the fluid.

Fluid friction is very much less than the sliding friction between solid surfaces.

THE SPECIAL CASE OF FREE ROLLING

When a wheel rolls along a surface without skidding, there is no relative motion between the surface and the wheel at the point of contact. Accordingly, there is no frictional force at the point of contact. If a force is applied to accelerate the wheel, that force will tend to cause skidding, that is, relative motion between the wheel and the surface at the point of contact. There is then a frictional force, just as in the case of sliding at the beginning of this note.

This frictional force provides a torque which increases the rotational speed of the wheel so that the skidding stops, and the frictional force goes away.

OVERCOMING FRICTION

There are three common ways of overcoming sliding friction.

We may change the nature of the contacting surfaces.

> This can be done by polishing or by applying a layer of low friction material, such as Teflon[7], to one surface. Some chemicals can also change the frictional characteristics of particular surfaces. Molybdenum disulphide is such a chemical.

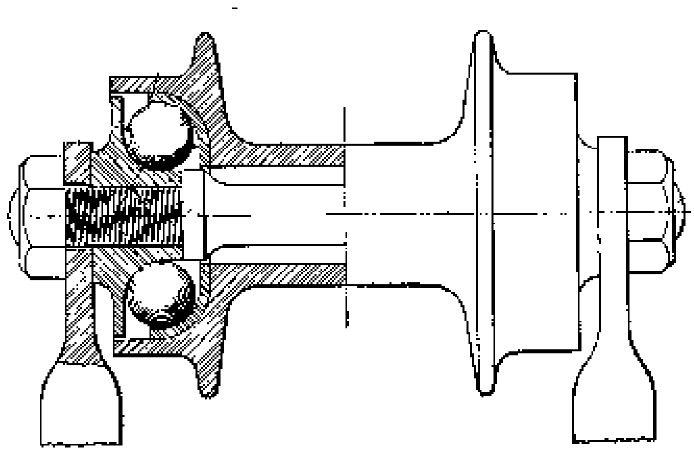
We may arrange for the surfaces to be kept slightly apart by a film of fluid such as an oil or a grease.

> This is called lubrication.

We might install cylindrical or spherical rollers between the contact surfaces.

---

[7] polytetrafluoroethylene



Do you recognise this friction reducing device?

VISCOSITY

Viscosity is the term used to describe the tendency of a fluid to stick to itself.

Viscosity is exhibited by both gases and liquids. Water, alcohol, oxygen and steam exhibit viscosity, as do mixtures such as oils and air.

Viscosity arises from the molecular nature of the fluid. Honey has quite a high viscosity, olive oil has rather less, and water is considerably less viscous than both of those liquid

You will be aware that honey, the more viscous liquid, pours much less readily than water. This arises in the following way. Because of molecular attraction and some mechanical interlocking, fluids stick to the solid surfaces. They stick to the walls of pipes and containers and they stick to plates immersed in the fluid - so much so that engineers frequently analyse fluid motion in pipes and over surfaces as though the fluid molecules touching the solid surface are firmly adhered to it.

Now if the fluid in a pipe or container is moving in bulk, while the outside edges of the fluid are stuck to the container, some severe distortion of the fluid must be occurring. Viscosity, the tendency of the fluid to stick to itself, determines the resistance to such distortion. So viscosity determines the resistance to pouring.

Viscosity also contributes to the difficulty of forcing liquids through pipes. A volume of highly viscous fluid such as honey requires a lot more pressure to force it through a pipe in given time than a less viscous liquid such as water.

If we look at the converse of this description of fluid flow, we can see that if, say, a thin solid flat plate moves in a viscous liquid there will be a resistance to the motion and that the amount of resistance will depend upon the viscosity of the fluid.

The viscosity of a fluid is generally not the same in all conditions. In particular the viscosity of a liquid decreases as the temperature of the liquid increases. Honey is much easier to pour if it is warm. Some oils change viscosity about one hundred-fold as their temperature changes from -20°C to +20°C.

Conversely, the viscosity of a gas increases as the temperature of the gas increases. The viscosity of air increases 27% between 0°C and 100°C.

In some liquids the viscosity also changes with the rate of flow. Blood is one such liquid. The viscosity of blood appears to decrease as flow rate increases because blood is a suspension of platelets which align with the flow at higher flow rates.

Scientists and engineers express viscosity in two different way. The difference is purely a matter of algebraic convenience for the occasion. The dynamic viscosity of a fluid is the force exerted on each of two plates, one metre apart and each of one square metre surface area, between which the fluid flows at a speed of one metre/second. Dynamic viscosity is expressed in units of Newton second per metre squared, Ns/m2 Sometimes the old unit the “poise” will be seen along with its sub multiples. 1 Ns/m^2 = 10 poise

. Kinematic viscosity is dynamic viscosity divided by the density of the fluid. Kinematic viscosity is expressed in metres squared per second.

As an exercise, plot a graph showing the viscosity of two different fluids at temperatures between say 0°C and 100°C. Castor oil and water would be good choices. Use the library to obtain the data. For comparison, find out the viscosity of air at about 20°C and write the value on the graph.

Viscosity and Lubrication

Machinery is lubricated to reduce friction and wear between parts which rub together and in many cases to carry away heat generated by the rubbing. Oils and greases lubricate parts by developing and maintaining a very thin film of comparatively soft material between the rubbing parts. Frequently it is only possible to maintain a partial film. Sometimes gases are used to provide this film.

The film is often maintained by keeping the fluid under pressure.[8] It is easier to generate and maintain this pressure in the gap between moving parts if the lubricant has a high viscosity. On the other hand, a high viscosity lubricant can make the machine stiffer to move than a low viscosity lubricant. It is quite common for lubricant to leak away from the rubbing surfaces, that is, away from the “bearing”. This means that a flow of lubricant into the bearing must be maintained. A low viscosity lubricant will leak away more easily. .A compromise must be struck in the choice of viscosity for the lubricant in a particular application..

Viscosity is not the only determinant of the friction reducing qualities of a fluid. Engineers also identify a quantity they call unctuousness which describes the film forming properties and film tenacity of the lubricant. Unctuousness is important when pressure cannot be maintained in the lubricating fluid.

Engine Oils

The motor car engine represents quite a severe test for lubrication technology. While this is a highly specialised business, there are three aspects which can be touched upon here to give you a general idea of the very special nature of the product you are buying for your motor car, often without a second thought..

---

8 The pressure may be externally applied by a pump, or may be generated within the bearing itself by a hydrodynamic wedging action.

The operating temperature within the engine varies from about -10°C to about 100°C. - from a cold start in Stanthorpe to a very hot engine stuttering up Cunningham's Gap behind a semi trailer. The cold engine needs oil of initially low viscosity so that it can quickly be flooded with lubricant and easily cranked over by the starter motor. The hot engine requires an oil of initially high viscosity so that when the viscosity of the oil is lowered by temperature it is still adequate for pressure lubrication.

Automotive engineers have a grading system for denoting the viscosity of oils. The definition of these grades and the method of measuring them was specified by the automotive standards organisation of the Society of Automotive Engineers a long time ago.

A motor car engine likes to operate with oil of viscosity corresponding to about 30 grade. The modern "multigrade" engine oil has been so formulated that it is almost viscostatic. When a 20W - 50 oil gets hot, it behaves as though it was initially a high viscosity 50 grade oil and its viscosity falls to about 30 grade. When it is cold it behaves as though it had been a 20 grade oil and its viscosity increases to about 30 grade.

Engine oil must be durable. It must retain its lubricating qualities in the presence of combustion by-products, severe compressive and shearing loads, high temperatures, and dilution by small amounts of water. The oil must not dry out or evaporate.

Engine oil must have the ability to retain wear debris in suspension until the oil reaches the filter that can safely remove and store the debris. This "detergent" capability is a disadvantage in oils for unfiltered applications such as gear boxes.

New engines have quite different requirements than worn engines. Oils suitable for one are not always suitable for the other.

The automotive standards organisations specify fitness-for-purpose requirements and test procedures for all these aspects of oils. The performance of the oil is denoted by the designation of the standard with which it conforms. That designation is marked on the container in which you buy the oil. The American Petroleum Industry standards designated SF and SG and CC are typical modern specifications.

VISCOSITY AND FLUID FLOW (ADVANCED NOTES)

The influence of viscosity in fluid flow is far from simple.

Viscosity was unknown in the 1700s when the classical equations of dynamics were being written down by the now legendary mathematicians. Technically and mathematically perfect equations of inviscid, incompressible, steady fluid flow were devised in the 1700s. However, they predicted that the drag force on a body immersed in and moving relative to a fluid would always be zero. Of course, that was directly contrary to experience. D'Alembert labelled this a paradox in his "Opuscules mathematiques" of 1768. The search for an answer occupied many eminent mathematicians for 150 years. Ultimately the answer was found to involve the interrelated effects of viscosity and unsteadiness. (In a fluid flow, when the flow pattern at any point is not permanently the same, the flow is called "unsteady" or "non-stationary".) The analytically rigorous result is expressed in the "Navier Stokes" differential equations of classical fluid dynamics

Unsteady or non stationary flows are no place for the faint-hearted. They have been particularly difficult for aerodynamicists to deal with. That is perhaps mainly because they are not amenable to direct experimental measurement.

Indeed classical fluid dynamics is so difficult that there are very, very few situations in which analytical solutions to the Navier Stokes equations can be derived.[9]. Because of these difficulties, the estimation of fluid dynamic forces has always rested on experimental measurement and approximate mathematical modelling. Usually there is some attempt to verify that the model is consistent with the classical theory. A model is held to be applicable across a general class of flow patterns.

In 1882/83, Osbourne Reynolds devised a parameter intended as a predictor of the similarity of flow patterns in experiments at different scales. It is a measure of the ratio of inertia forces to viscous forces and a highly important modelling and force estimation tool.

$$\mathrm{Re} = \frac{Vl\rho}{\mu}$$

where:
V is the velocity of the fluid
l is a characteristic length
ρ is the fluid density and
μ is the fluid dynamic viscosity

The proposition is that a mathematical model will remain valid so long as Reynold's number is constant.

We should observe now that classical fluid dynamics does not apply at very low pressures. That is, at densities where the distance between molecules of fluid is of the same order as the size of objects of interest immersed in the fluid. - the molecular flow regime[10]. Classical fluid dynamics deals with an elastic continuum - rather than the momentum of individual molecules. For a similar reason, one might argue that Re is not defined at such low pressures.

To illustrate the complexity of the influence of viscosity on fluid flow let us look at some concepts in aerodynamic drag. The drag is classified into two sorts according to cause. We start by visualising air flowing past a stationary body.[11]

The first sort of drag is due to viscosity or the tendency of the air to stick to itself and other substances. This we usually call skin friction. When we think of the air moving over the body, viscous effects lead to a loss of momentum, and perhaps microscopic turbulence in a thin sheet or tube of wake. The drag force on the body corresponds to the loss of momentum from the air.

The second sort is caused by "separation" of the fluid from the surface of the body. The inertia of the fluid overcomes the tendency of the viscous fluid to stick to the body. There is sensible turbulence about part

---

9 But in recent times they have been used as the basis of flow modelling by finite element techniques in high speed computers.

10 At room temp and 760mm Hg pressure the mean free path of air molecules is about 10^{-4}mm. At 1mm Hg the mean free path is about 10^{-1}mm.

11 With aircraft, vehicles and projectiles of course, the air doesn't move over a stationary body - the body moves through the air. This is an important but not well recognised distinction. In these cases, the body pushes on the air and increases its momentum. A dynamic equivalence between stationary and moving air can be established for some sorts of flow. This is done mathematically by the Euler transformation which Euler's contemporary, the mathematician d'Alambert is likely to have had something to say about circa 1770.

of the body and in the wake; and a loss of wake momentum in the direction of the general stream of air. Again the drag force on the body corresponds to the loss of momentum. This is the sort of thing that happens when the air is asked to turn too sharply - at a sharp edge for example. When separation occurs, it causes more drag than viscosity would cause.

At Reynolds Numbers above 50 to 100, the dynamic forces so exceed the viscous forces that some flow separation happens for all blunt bodies and sharp edges. Eddies develop near the trailing edge of the body. Vortexes are shed into the wake alternately from one side of the body and the other. Once well developed, the wake adopts a mathematically describable form called a "Kármán vortex street" after Theodore von Kármán. The frequency of shedding increases steadily with Reynolds Number. It makes telegraph wires, and the strings of King David's harp, sing in the wind.

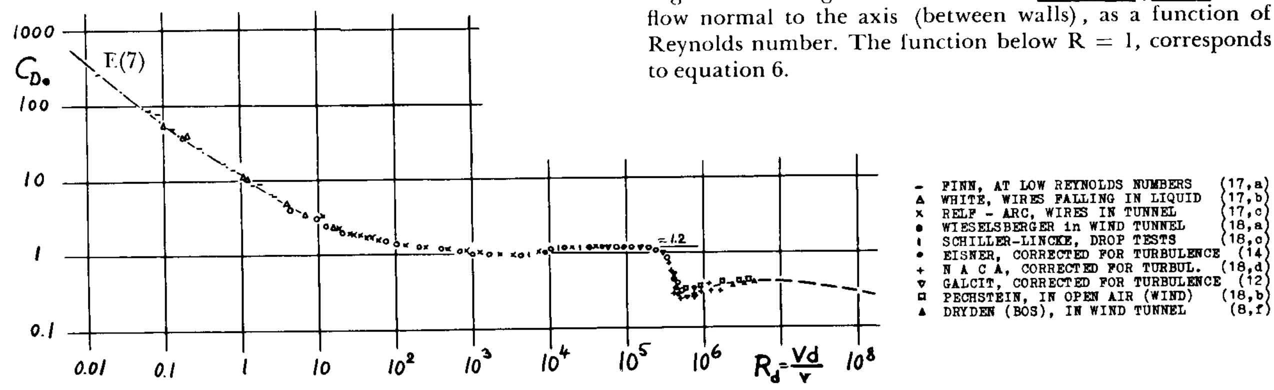
To make things easy for themselves, aerodynamicists invent a boundary layer close to the surface of the body and say that all the viscous effects take place in that layer. If this boundary layer flows in a laminar fashion, one where there is no transfer of energy across the thickness from one layer to another, the skin friction will be low but there will be a tendency for the flow to separate easily. If the boundary layer is turbulent, so that energy is transferred across its thickness by the turbulence, skin friction will appear to be greater, but there will be less tendency for the flow to separate. So a turbulent boundary layer may result in less drag over the range of circumstances the body experiences.

Boundary layers are at first laminar, but they change from laminar to turbulent. The change happens at a Reynolds Number[12]. of about 10^6 . Increasing airspeed or the size of the immersed body will do the trick. The transition needs to be triggered and often this happens because of roughness on the body or turbulence in the free stream. Under some circumstances, the transition from a laminar to turbulent boundary layer can delay separation, thereby reducing the drag. This is a very well known phenomenon. It is not confined to cylinders or spheres, though they are the best documented shapes. It is less likely to happen if separation is being triggered by a sharp edge irrespective of the state of the boundary layer.

The following graph has been taken from the industry bible "Fluid Dynamic Drag" by F.S Hoerner. It shows the variation of drag coefficient with Reynolds Number for cylinders without ends. (Drag = $\frac{1}{2}\rho V^2 A C_D$)

12 This Reynolds Number is a property not only of the air stream, but also of the stream-wise position of the observation point on the body - the distance back from the leading edge.

Figure 12. Drag coefficient of the circular cylinder in a flow normal to the axis (between walls), as a function of Reynolds number. The function below R = 1, corresponds to equation 6.



This is one of the most famous graphs in aerodynamics. It represents the work of hundreds of experiments. As Re changes there are several changes in the pattern of fluid flow around a cylinder.

At low values of Re <1, we have one of the few cases where drag is accurately predicted by classical fluid dynamics. Here the flow is smooth and unseparated, but, the flow behind the cylinder has an appreciable change in speed. By Re=10, two symmetrical standing vortices form behind the cylinder. The vortices elongate as Re increases. When Re reaches 40 or 50, the vortices become asymmetrical. They detach from the cylinder and move downstream as if they were being shed alternately from on side of the cylinder then the other. An eddying motion is set up in the wake which further downstream looks like a generalised random turbulence. For Re between 50 and 150 the shedding of vortices is regular. The eddying motion is periodic both in time and place, and the flow can be described by the Kármán vortex street. The vortex shedding becomes irregular between Re = 150 and 300. More and more small scale turbulence appears in the wake. Then, at Re = $5x10^5$ we see the boundary layer transition effect already described and the vortex pattern disappears altogether n the general turbulence..

The relevance of all that to our topic is that Reynold's number is a function of fluid viscosity. So that the occurrence of each of the changes in flow patterns is viscosity dependent.

The influence of viscosity on fluid flow is far from simple.

Air is a mixture of gases including steam. Steam is very viscous stuff. You may care to speculate on the effect of hot and humid conditions on the flight of a cricket ball. Will the viscosity of the air be higher than in cool dry conditions? What does that mean for Reynolds Number? What does that mean for the point at which the airflow separates from the surface of the ball?

ESTIMATION AND EXPRESSION OF ERRORS

In reporting the results of any experiment, no matter how trivial, we must always endeavour to give the recipient of the report an idea of how inaccurate our results might be or how much confidence we have in them. In the past, the uncertainty about results has been loosely termed "error" but that is not strictly logical.

We may present a general discussion of the sources of uncertainty, or we may sometimes present the uncertainty numerically. It is usual to have a numerical presentation when the experiment consists mainly of direct measurements and deductions from those measurements.

LIMIT OF ESTIMATION

Whatever kind of instrument is being used to make a measurement, there will always be some uncertainty because of the way the instrument presents its results to us. No instrument presents a result with infinite precision.

This particular limitation to the accuracy is known as the limit of estimation of the reading.

For a digital instrument the limit of estimation is normally stated to be ±1 in the last digit on the instrument's readout.

For an analogue instrument the limit of estimation will depend as much on the observer's ability to estimate subdivisions of scale intervals as on the scale itself.

As an example, consider the measurement of the length of an object using a rule with a scale graduated in equal divisions. Suppose that one end of the object is placed against the zero mark of the scale and that the other end is found to be nearest to the nth scale mark. Then the length lies between n - ½ and n + ½ scale divisions and no matter how many times the measurement is repeated the limit of estimation will always be the same. It is conventional to express the result as n ± ½ scale divisions eg 51 ± ½ mm. We say the limit of estimation is ± ½ the scale division.

It has been assumed here that we can only read to the nearest scale division, but in practice it is usually possible to read to a smaller interval by visually estimating subdivisions of the scale. So we have a "limit of reading" smaller than a scale division. In these circumstances we say the limit of estimation is ± ½ the limit of reading

The reading is always reported to a correspondingly limited number of significant figures. It would not be normal to report that, for example, the length of an object was 10.006mm ± 0.05m

Because the limit of estimation is always the same no matter how many readings are taken, we say it is a "systematic" uncertainty, to distinguish it from errors which change randomly from one measurement to the next.

If, because of the possible existence of random variations, we measured something ten times and took the mean as the best answer, we would still have to say that there was a systematic uncertainty in that answer.

That uncertainty would still be the limit of estimation due to the instrument we used.

The limit of reading is not the only source of uncertainty in instruments, particularly complex instruments, but including instruments so apparently simple as the moving coil ammeter or voltmeter.

An Example

Suppose we have a rectangular piece of sheet metal and we measure the length of each of its sides, A B, C, D with a steel ruler. If the ruler is graduated in millimetres and we think we can read it to the nearest half millimetre, we could write our answers thus:

A = 10.5 ±0.25 mm

B = 10.0 ±0.25 mm

C = 5.5±0.25 mm

D = 5.5 ±0.25 mm

- The term ±0.25 mm is often called the "absolute" uncertainty. We can also express the uncertainty as "proportional" uncertainty. The proportional uncertainty is the absolute uncertainty divided by the estimated value. The proportional uncertainty in A is ±0.25 / 10.5 that is 1/42, more usually expressed as 2.4%. In our example the estimates and proportional uncertainties are :

A = 10.5mm ± 2.4%

B = 10.0mm ± 2.5 %

C = 5.5mm ± 4.5%

D = 5.5mm ± 4.5%

We can now go on to estimate the perimeter and the area of our rectangle. The perimeter is equal to the sum of the lengths of the sides = A+B+C+D , So:

Perimeter = 31.5mm.

The area is equal to the length of the rectangle multiplied by the width. But, we have two different values for the length, A and B. One way to proceed is to use the mean of A and B as the length. We would then say that the best estimate of the length is (A+B)/2 = 10.25mm. The absolute uncertainty in this value is still half the limit of reading, that is, ±0.25mm. The proportional uncertainty in the length is 0.25/10.25 = 2.4%. Thus we have :

Length = 10.25mm ±2.4%

Width = 5.5mm ± 4.5%

Area = 56.4 mm^2

Combination of Uncertainties

Having now determined the perimeter and area of the rectangle. How do we express the uncertainty of our estimates

. Because calculating the perimeter involves *adding* the lengths if the sides, we express the *absolute* uncertainty in the estimate of the perimeter by *adding the absolute uncertainty* in the measurement of the sides. Thus we express the perimeter as 31.5 ± 1mm.

Now, because calculating the area involves *multiplying* the length by the width, we express the *proportional* uncertainty in the estimate of the area by *adding* the *proportional uncertainties* in the measurement of the length and width of the rectangle. Thus the proportional uncertainty in the area is 2.4% + 4.5% = 6.9% which we would round to 7%. Lastly we note that 6.9% of 56.4 mm^2 is 3.9 mm^2 . So we would report that the either that the area was 56 mm^2 ±7% or that the area was 56 ± 4 mm^2 .

ADVANCED NOTES

. Combination of Uncertainties

If the value F of a physical quantity depends on a number of dimensions x,y,z,… so that

F = f(x,y,z,..)

then, if the measurements are all independent, the component of the systematic uncertainty in F associated with the systematic uncertainty dx in x is given by :

$$dF_x = \left| \frac{\partial F}{\partial x} \right| dx$$

There is no rigorous method of combining the various dF to obtain the overall systematic uncertainty. One method that has been used combines them by arithmetic addition, in the manner of the calculus.

$$dF = \left|\frac{\partial F}{\partial x}\right| dx + \left|\frac{\partial F}{\partial y}\right| dy + \left|\frac{\partial F}{\partial z}\right| dz + \ldots$$

This method is likely to overestimate the size of the total systematic uncertainty and can be considered as an estimate of the maximum possible limit.

In the simple case when F= x+ y+ z this method says that the total uncertainty is the sum of the individual uncertainties.

When F = xyz this method says that the total uncertainty, expressed as a fraction, is the sum of the individual fractional uncertainties. That is:

$$\frac{dF}{F} = \frac{dx}{x} + \frac{dy}{y} + \frac{dz}{z} + \ldots$$

Randomization

Although the limit of estimation is usually assessed by such nonstatistical means and is thus treated as a systematic uncertainty, this need not necessarily be the case. The limit of estimation of an instrument is an example of an uncertainty that is usually systematic. but which can, by change of experimental procedure, be replaced by a random uncertainty. Because of the importance of the underlying concept the example will be discussed in more detail.

It is possible to randomize the limit of estimation by placing one end of the object, not at a fixed zero, but at points chosen at random near the lower end of the scale. Suppose the object has a true length n+x scale intervals where x is less than one scale interval. When the object is placed at an arbitrary position on the scale it will be found to overlap either n or n + 1 scale divisions, the probability of a reading n being (1- x) and the probability of a reading n + 1 being x. If R separate measurements are made and the object is placed at random on the scale each time, and, if r of the readings have a value n, then R - r have a value n + 1, and the ratio (R - r)/R gives an unbiased estimate of x. In other words :

length of object = $n+\dfrac{R-r}{R}$ scale divisions.

Furthermore, the variance of the mean value from R measurements is given by

$$S^2\,(n+x) \approx \frac{r(R-r)}{R^3} \approx \frac{x(1-x)}{R}$$

That is, the standard error of the mean is inversely proportional to √R for large R. (See JEFFREYs, H. Scientific Inference, 2nd edition, Cambridge University Press, 1957, 61.) Thus the precision attainable is limited only by the number of observations made, and hence the uncertainty becomes a component of the random uncertainty. If there are other additional causes of random variations in the readings then the value of the standard error of the mean will automatically include all such components. It should be noted that while the reduction in the random uncertainty is limited only by the number of measurements taken, in principle at least there will be residual systematic uncertainties due, for example, to the scale non-linearity of the ruler. For other analogue instruments the randomization can be achieved by resetting the zero of the instrument randomly at points near the lower end of the scale before each measurement is made. The concept is also applicable to digital instruments where, in some cases, the randomization is inherent in their operation.

Thus the above is an example of how, by a change of procedure, a subjectively assessed systematic uncertainty can be reduced at the expense of introducing a random uncertainty whose magnitude depends on the number of observations. In examining the accuracy that might be achieved in an experiment it is always worthwhile to consider whether the major components of the systematic uncertainty can be reduced by adopting randomization techniques similar to that described above.

In deciding whether an uncertainty is to be regarded as random or systematic, the criterion should be whether the values assigned to the uncertainty limits were derived from a statistical analysis of a number of measurements (assuming a Normal or other appropriate distribution), or whether it was possible only to estimate limiting values by a nonstatistical assessment. It follows from this recommendation that, once they have been classified as random or systematic, the uncertainties on a result should remain in their original category regardless of how that result is subsequently used. The two categories should be kept separate in the progressive combination of uncertainties throughout a complex experiment, in the statement of the final result of that experiment, and when using that result in any further experiments. Finally it is recommended that randomization techniques be introduced wherever practicable into experimental procedures in order to reduce the systematic uncertainties.

With acknowledgements to National Physics Laboratory