It's about Time

Time is nature's way of preventing everything happening at once.

John Wheeler

Let us first discuss how astronomers measured the passage of time until the 1960's.

Local Solar Time

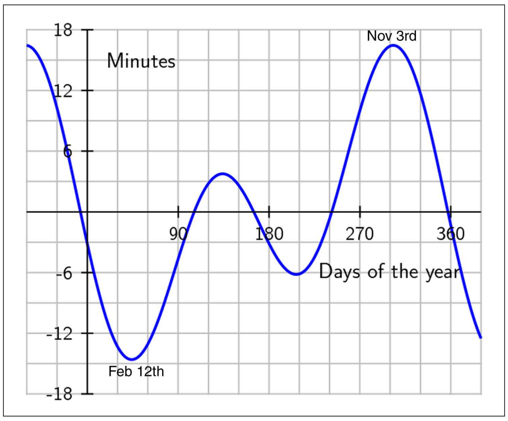
For centuries, the time of day was directly linked to the Sun's passage across the sky, with 24 hours being the time between one transit of the Sun across the meridian (the line across the sky from north to south) and that on the following day. This time standard is called "Local Solar Time" and is the time indicated on a sundial. The time such clocks would show would thus vary across the United Kingdom, as Noon is later in the west. It is surprising the difference this makes. In total, the United Kingdom stretches 9.55 degrees in longitude from Lowestoft in the east to Mangor Beg in County Fermanagh, Northern Ireland in the west. As 15 degrees is equivalent to 1 hour, this is a time difference of just over 38 minutes!

Greenwich Mean Time

As the railways progressed across the UK, this difference became an embarrassment and so London or "Greenwich" time was applied across the whole of the UK. A further problem had become apparent as clocks became more accurate: due to the fact that, as the Earth's orbit is elliptical, the length of the day varies slightly. Thus 24 hours, as measured by clocks, was defined to be the *average* length of the day over one year. This time standard became known as **Greenwich Mean Time** (GMT).

The Equation of Time

The use of GMT has the consequence that, during the year, our clocks get in and out of step with the Sun. The difference between GMT and the local solar time at Greenwich is called the "**Equation of Time**". The result is that the Sun is not always due south at noon - even in London - and the Sun can transit (cross the meridian) up to 16 minutes 33 seconds before noon as measured by a clock giving GMT and up to 14 minutes 6 seconds afterwards. This means that sunrise and sunset are not usually symmetrically centred on midday and this does give a noticeable effect around Christmas time Though the shortest day is on December 21st, the Winter Solstice, the earliest sunset is around December 10th and the latest sunrise does not occur until Jan 2nd, so the mornings continue to get darker for a couple of weeks after December 21st whilst, by the beginning of January, the evenings are appreciably longer.



The "Equation of Time" - the difference between GMT and local solar time at Greenwich Observatory.

Universal Time

Greenwich Mean Time was formally replaced by **Universal Time** (UT) in 1928 (though the title has not yet come into common usage) but was essentially the same as GMT until 1967 when the definition of the second was changed! Prior to this, one second was defined as one 86,400th of a mean day as determined by the rotation of the Earth. The rotation rate of the Earth was thus our fundamental time standard. The problem with this definition is that, due to the tidal forces of the Moon, the Earths rotation rate is gradually slowing and, as a consequence, the length of time defined by the second was increasing! So, in 1967, a new definition of the second was made:

> **The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom.**

Thus our clocks are now related to an Atomic Time standard which uses Caesium Beam frequency standards to determine the length of the second.

But this has not stopped the Earth's rotation from slowing down, and so very gradually the synchronization between the Sun's position in the sky and our clocks will be lost. To overcome this, when the difference between the time measured by the atomic clocks and the Sun (as determined by the Earth's rotation rate) differs by around a second, a leap second is inserted to bring solar and atomic time back in step. This is usually done at midnight on New Year's Eve or the 30th June. Since the time definition was changed, 22 leap seconds have had to be added, about one every 18 months, but there were none between 1998 and 2005 showing the slowdown is not particularly regular. Leap seconds are somewhat of a nuisance for systems such as the Global Positioning System (GPS) Network and there is pressure to do away with them which is, not surprisingly, opposed by astronomers! If no correction were made and the average slow down over the last 39 years of 0.56 of a second per year continues, then in 1000 years UT and solar time would have drifted apart by ~ 9 minutes.

Sidereal Time

If one started an electronic stop watch running on UT as the star Rigel, in Orion, was seen to cross the meridian and stopped it the following night when it again crossed the meridian, it would be found to read 23 hours, 56 minutes and 4.09 seconds, not 24 hours. This period is called the **sidereal day** and is the length of the day as measured with respect to the apparent rotation of the stars.

Why does the sidereal day have this value? Imagine that the Earth was not rotating around its axis and we could observe from the dark side of the Earth facing away from the Sun. At some point in time we would see the star Rigel (in Orion) due south. As the Earth moves around the Sun, Rigel would be seen to move towards the west and, three months later, would set from view. Six months later after setting in the west, it would be seen to rise in the east and precisely one year later we would see it due south again. So, in the absence of the Earth's rotation, Rigel would appear to make one rotation of the Earth in one year and so the sidereal day would be one Earth year. But, in reality, during this time, the Earth has made ~365 rotations so, in relation to the star Rigel (or any other star), the Earth has made a total of ~ 365 +1 rotations in one year and hence there are ~366 sidereal days in one year. The sidereal day is thus a little shorter and is approximately 365/366 of an Earth day.

The difference would be ~ 1/366 of a day or 1440/366 minutes giving 3.93 minutes or 3 minutes 55.8 seconds. The length of the sidereal day on this simplified calculation is thus approximately 23 hrs 56 minutes 4.2 seconds, very close to the actual value.

Clocks

Sundials

These are perhaps the most fundamental clocks of all - they keep, by definition, local solar time. They are not, unfortunately, as useful in the UK as more southerly countries! They are made in many forms; some with a horizontal flat face with gnomon pointing up to the pole star, some with vertical faces on the sides of buildings and some where a band stretches around the gnomon. There is quite an art, and some mathematics, employed in their design with a particularly interesting one, where a person acts as the vertical gnomon, called an analemmatic sundial.

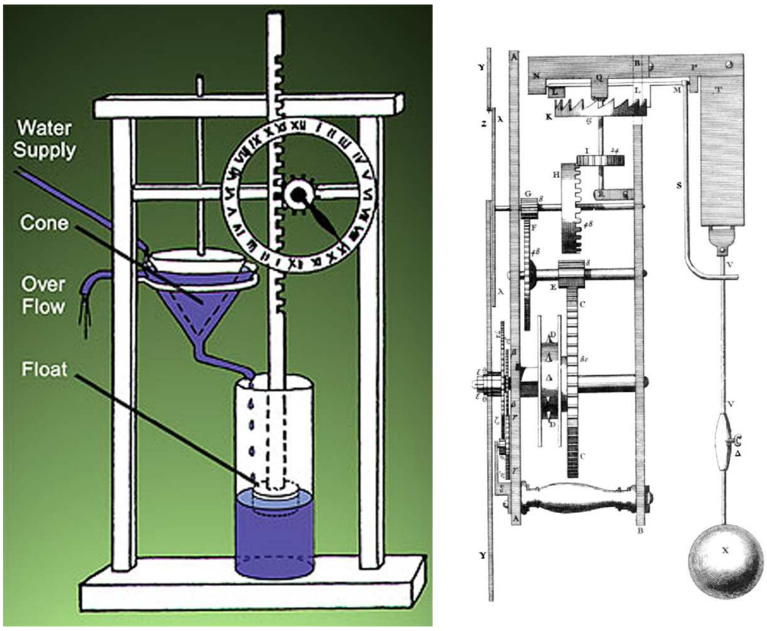
Water Clocks

These are rather fun, and until the invention of the pendulum clock, the most accurate. The simplest form just filled a cylinder of constant cross section from a steady supply of water. As in all water clocks, the key to accuracy was to have a constant "head" of water so that the water flow into the clock was constant. A good way the achieve this was to allow more water that required for the clock itself to pass into a reservoir which thus continuously overflows and so is kept brim full. An exit pipe at some distance below the surface will thus have a constant head of water above it and so be at a fixed pressure. Later, water clocks were designed to provide a mechanical system to move an hour hand (and perhaps a minute hand) around a dial, for example in the simple cylinder design, a float might rise carrying a toothed vertical arm mounted on it to rotate the hour hand of a clock face. Later, some clocks were based on a water wheel which would rotate at a constant speed to drive the hands through a series of gears.

Pendulum Clocks

The idea of using a pendulum to keep time is attributed to Galileo who, as a student in 1602, is said to have watched a suspended lamp swing back and forth in the cathedral of Pisa. Galileo's discovery was that the period of swing of a pendulum (at least for relatively small swings) is independent of its amplitude - the so called “isochronism” of the pendulum. In 1603 a friend of his began to use a short pendulum to measure the pulse rate of his patients.

In 1641, at the age of 77 and totally blind, Galileo, aided by his son, turned his attention to using a pendulum to construct a clock but although drawings were made and a clock partly constructed in 1649, it was never completed. Galileo’s work inspired the Dutch scientist Christiaan Huygens in 1657 to invent and patent a working pendulum clock. His first design used a verge escapement, which required quite a wide pendulum swing causing its period to be somewhat variable. He made a second design which used gears to limit the swing and then, in his third design, used curved “jaws” to effectively change the length of the pendulum dependant on the swing so correcting the problem to a large extent.



A Water Clock and Christiaan Huygen's first pendulum clock.

Later, in 1670, clockmakers invented the anchor escapement which reduced the pendulum's swing to 4°-6°. This allowed the clock's case to accommodate longer, slower pendulums - in particular, the “seconds” pendulum (also called the Royal pendulum). The length of the pendulum is about 1 metre with each swing taking one second. The tall narrow clocks built using these pendulums became known as grandfather clocks and, because of their increased accuracy, a minute hand began to be added after 1690.

One problem was that pendulum clocks were observed to slow down in summer due to the thermal expansion of the pendulum rod. This was solved by the invention of the mercury pendulum, which had a mercury vessel as its bob, and the gridiron pendulum that used alternating rods of iron and zinc. More recently, pendulums made of Invar, a steel–nickel alloy, having an exceedingly low coefficient of expansion have been used. A key objective to obtain high precision is to try to allow the pendulum to run as freely as possible and the most accurate pendulum clocks (called regulators) only give a sustaining pulse to the pendulum every 30 seconds.

The pendulum swing (and hence period) will also be slightly affected by changes in the barometric pressure. A bellow device which changes its size as a function of pressure can be used to compensate, but some of the very best regulator pendulums are operated in a near vacuum. It should be noted that, as the period of a pendulum is a function of the gravitational pull of the Earth, they have to be calibrated for both their height above sea level and the latitude of its location! (The effective gravitational pull of the Earth at the equator is reduced relative to the poles due to the Earth's rotation.)

For many years, regulators, located in observatories to allow astronomical calibration, served as the primary standards for national time distribution services. Initially, the US time standard used Riefler pendulum clocks, accurate to about 10 milliseconds per day. In 1929 it switched to the Shortt free pendulum clock (about 1 second per year) before phasing in quartz time standards in the 1930s.

Quartz Clocks

A quartz clock uses an electronic oscillator that is regulated by a quartz crystal to keep time. They are at least an order of magnitude more accurate than good mechanical clocks. In most modern quartz clocks or watches, the quartz crystal resonator is in the shape of a small tuning fork, laser-trimmed to vibrate at 32,768 Hz. This frequency is equal to 2^{15} Hz. A power of 2 is chosen so a chain of 15 digital divide-by-2 stages can derive the 1 Hz signal which then drives the clock or watch's second hand. A typical quartz wristwatch will gain or lose less than a half second per day at body temperature.

If a quartz watch is kept at a reasonably constant temperature it can be accurate within 10 seconds per year. To improve accuracy, quartz chronometers which are to be used as time standards include a crystal oven to keep the crystal at a constant temperature. From the 1930's, quartz time standards replaced pendulum regulators in providing national time standards. In 1932 a quartz clock was able to measure the tiny weekly variations in the rotation rate of the Earth and that (due to the Moon's tidal forces) the rotation rate was slowing down! As a second was then defined as 1/84,600th of a day, this of course means that the period of a second was not constant – not accepted kindly by physicists!

Atomic Clocks

Quartz time standards remained in use until the 1960's when they were replaced by atomic clocks. These are the most accurate time and frequency standards known, and use the precise microwave signal that electrons emit or absorb when they change energy levels in an atom. They provide accuracies of approximately 1 part in 10^{14} which is ~ 10^{-9} seconds per day!

The first accurate atomic clock was built by Louis Essen in 1955 at the National Physical Laboratory in the UK and used a beam of caesium-133 atoms passing through a cylinder which acts as a resonant cavity at the frequency emitted by the caesium atoms. Such caesium beam clocks provide the fundamental time standards of most nations, but are very expensive and usually backed up with Hydrogen Maser atomic clocks such as that at Jodrell Bank.

The hydrogen maser uses the fact that, when in a magnetic field, the lowest energy level of hydrogen is split into two. In the upper energy level, the spins of the proton and electron are parallel whilst in the lower, anti-parallel. A beam of hydrogen atoms is produced (having equal numbers in both states) which is passed through a special (hexapole) magnet which splits them into two beams. The beam of hydrogen atoms in the higher state is passed into a resonant cavity which contains radiation at the frequency corresponding to the transition from the upper to the lower state - 1,420,405,752 Hz. This radiation stimulates the arriving atoms to radiate and build up the level of radiation in the cavity. A small probe extracts a small amount of energy from the cavity which is used to lock a crystal oscillator to a frequency with equal precision. This frequency can then be divided down to give a "pulse" at 1Hz to drive a clock.

The resonant cavity of a Hydrogen Maser.

The most common atomic clocks use excited rubidium atoms. They are inexpensive but are inherently less accurate. However, they can be periodically corrected by a GPS receiver to achieve long-term accuracy equal to the U.S. national time standards.

The most accurate atomic clock in continuous use today is NIST-F1, which is now the USA's primary time and frequency standard. It is a caesium fountain atomic clock which extracts the resonant frequency (9,192,631,770 Hz) of the caesium atoms when they are virtually stationary. Six infrared laser beams gently push the caesium atoms together into a ball which slows down the movement of the atoms and cools them to temperatures near absolute zero. This beautifully removes the effect of the Doppler shift that affects atomic clocks that use atoms in motion. The precision given by

NIST-F1 is now about 5×10^{-16}, which means it would neither gain nor lose a second in more than 60 million years and is about ten times more accurate than the caesium beam atomic clock that served as the United State's primary time and frequency standard from 1993-1999.

Radio Controlled "Atomic Clocks"

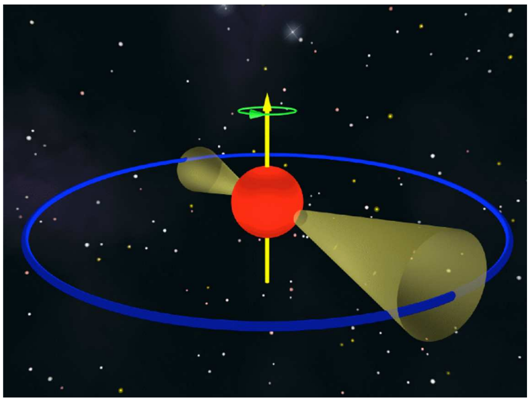
Such clocks and wristwatches are now widely available and are based on quartz watch movements but with additional circuitry to receive time signals from a number of longwave radio transmitters around the world such as "MSF" in the UK. These signals are used to correct the time displayed by the clock - often around midnight - and can even adjust for British Summer Time. They will normally be accurate to the second which is good enough for most people. An interesting point is that such a clock in London will be about 2 milliseconds slow as it takes this time for the time signal to reach London from the transmitter in Cumbria!

Pulsars - the best natural clocks in the universe

Pulsars were discovered serendipitously by Jocelyn Bell in 1973 when she discovered a radio source that was giving a series of very regularly spaced pulses – hence the name, pulsar, given them by the science correspondent of the Daily Telegraph. Fred Hoyle suggested that the signal might be pulsed emissions coming from an oscillating neutron star - the theoretical remnant of a supernova but never previously observed. Some three months later Thomas Gold at Cornell University in Ithaca, USA, gave a satisfying explanation for the pulsed signals.

Gold suggested that the radio signals were indeed coming from neutron stars, the remnants of giant stars, but that the neutron star was not oscillating, but instead spinning rapidly around its axis. He surmised that the rotation, coupled with the expected intense magnetic field generates two steady beams of radio waves along the axis of the magnetic field lines, one beam above the north magnetic pole and one above the south magnetic pole. If (as in the case of the Earth) the magnetic field axis is not aligned with the neutron star's rotation axis, these two beans would sweep around the sky rather like the beam from a lighthouse. If then, by chance, one of the two beams crossed our location in space, our radio telescopes would detect a sequence of regular pulses - just as Bell had observed - whose period was simply the rotation rate of the neutron star.

Gold, in this paper, pointed out that a neutron star (due to the conservation of angular momentum when it was formed) could easily be spinning at such rates. He expected that most pulsars should be spinning even faster than the first two observed by Jocelyn Bell and suggested a maximum rate of around 100 pulses per second.



Twin beams emitted by a Pulsar.

Since then, nearly 2000 pulsars have been discovered. The majority have periods between 0.25 and 2 seconds. It is thought that as the pulsar rotation rate slows the emission mechanism breaks down and the slowest pulsar detected has a period of 4.308 seconds.

Millisecond Pulsars

There is a class of "millisecond" pulsars where the proximity of a companion star has enabled the neutron star to "pull" material from the outer envelope of the adjacent star onto itself. This also transfers angular momentum so spinning the pulsar up to give periods in the millisecond range - hence their name. The fastest known pulsar is spinning at just over 700 times per second - with a point on its equator moving at 20% of the speed of light and close the point where it is thought theoretically that the neutron star would break up!

Pulsars slowly radiate energy, which is derived from their angular momentum. This is so high that the rate of slowdown is exceptionally slow and so pulsars make highly accurate clocks and some may even be able to challenge the accuracy of the best atomic clocks. One of the best pulsar clocks known at the present time is 1713+07 which has been "spun up" by matter falling onto it from a companion white dwarf star. It now has a pulse period of 4.57 milliseconds - spinning 218.8 times per second - and is currently slowing down at a rate of 200 nanoseconds in 12 years. That is a precision of one part in 1,892,160,000,000 ~ better than one part in 10^{13}!

An absolute time standard - Cosmic Time

In 1905, Albert Einstein, then working in the Berne Patent Office, published his paper on the Special Theory of Relativity. Perhaps one of the most well known aspects of this theory is that moving clocks appear to run slow when compared to a clock at rest with an observer - a phenomena called **time dilation.** This prediction has been proven by flying highly accurate atomic clocks around the world and has to be taken into account in the Global Positioning System (GPS) used for navigation.

As time is relative can we actually define a time standard with which to observe the evolution of the universe? One could, perhaps, define what might be called **cosmic time** as that measured by a clock that is stationary with respect to the universe as a whole. But how would this time relate to clocks on Earth? We know that the Earth is moving around the Sun, and that the Sun is moving around the centre of our Milky Way galaxy once every ~220 million years. But can we measure how fast the solar system is moving with respect to the universe?   Perhaps surprisingly, we can.

Since 1965, observations have been made of what is called the Cosmic Microwave Background (CMB) – radiation that originated near the time of its origin and which now pervades the whole universe. This radiation is very largely composed of a mix of long wavelength infra-red and very short wavelength radio waves - it has a "blackbody spectrum" that will be discussed in the next chapter. For simplicity, just suppose that it is made up of only one wavelength and that the solar system is moving in a certain direction with respect to this radiation. The Doppler effect will alter the apparent wavelength that we observe so that, when looking along the direction in which the solar system is moving it will be blue shifted and appear to have a shorter wavelength. Conversely, in the opposite direction, the radiation will appear to be red shifted and have a longer wavelength. From very precise measurements of the CMB we now know that we are moving through towards the constellation Leo at a speed of ~ 650 km/second. (2,340,000 km/hr, or about 0.22% of the speed of light!) This is thus our speed with respect to the universe as a whole.

We can thus calculate how the time of a clock at rest with the universe – measuring cosmic time – will differ from our clocks. To do this we need to derive the formula that determines the observed time dilation as a function of relative speed. This is not difficult if we can imagine a very simple "clock".

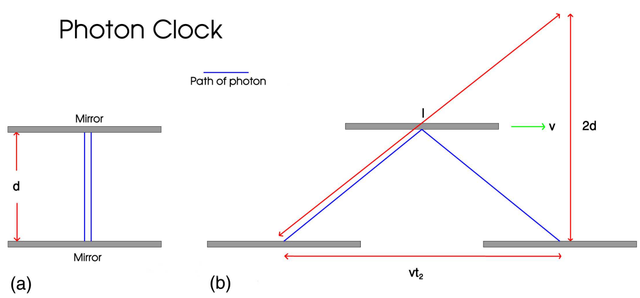


Diagram (a) show a photon clock at rest with the observer, whilst diagram (b) shows a photon clock moving at a speed v with respect to the observer.

The clock is made by reflecting a photon back and forth between a pair of perfect mirrors separated by a distance, d, as seen in the figure part (a). Our "tick" happens every time the photon reflects off the lower mirror and so the photon will travel a distance 2d between each tick. Our fundamental time period, t_1, will thus be given by:

$$t_1 = 2d / c$$

Suppose we observe such a clock moving past us at speed v. We will see the situation shown in part (b). As seen from our point of view, the photon will have to travel a longer distance, l, between each tick. This distance is given by

$$l = ((2d)^2 + (vt_2)^2)^{1/2}$$

So the time interval between each tick, t_2, will then be given by

$$t_2 = l / c = ((4d^2 + v^2 t_2^2) / c^2)^{1/2}$$

(c has been squared and put inside the square root.)
Squaring both sides and cross multiplying gives

$$t_2^2 c^2 = 4d^2 + v^2 t_2^2$$

We can now relate t_2 and t_1 to v by substituting for d from above using $d^2 = t_1^2 c^2 / 4$, giving,

$$t_2^2 c^2 = t_1^2 c^2 + v^2 t_2^2$$

and

$$t_2^2 (c^2 - v^2) = t_1^2 c^2$$

so, finally,

$$t_2 / t_1 = c^2 / \textbf{ sqrt } (c^2 - v^2)$$

or,

$$\boldsymbol{t_2 / t_1 = 1 / \text{ sqrt}(1 - v^2 / c^2)}$$

This is the time dilation formula, giving the ratio of time intervals as a function of the relative speed v. Note that the time dilation only become significant when v approaches the value of c.

We can now enter our speed with respect to the universe, 650 km/sec, into this equation and get the ratio 1.0000023. This is exceedingly small so, to a very good approximation, our clocks can be used to measure the time scale of the universe.

[Note: The effect of gravitational time dilation which is described below also needs to be considered. Due to this effect, clocks on the Earth's surface run slow compared to a clock in free space by ~700 picoseconds per second which is the order of 1 part in ~ 10^{-9}. This is a far smaller effect than caused by our passage through space and thus can be ignored.]

Muon Decay

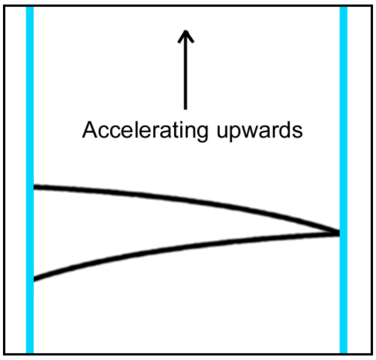
The time dilation predicted by Einstein's theory has been tested many times. Perhaps the simplest demonstration is given by the fact that we can observe muons at the surface of the Earth. Muons are radioactive particles which decay into an electron and 2 neutrinos with a half life of 1.56 microseconds (2×10^{-6} seconds) measured when *they are at rest*. This means that after 1.56 microseconds half will have decayed. Many are produced at a height of ~10 km in the upper atmosphere by the influx of cosmic rays and travel towards the ground at a speed of ~0.98c. They would thus take ~ 34 microseconds to reach the ground. 34 microseconds is nearly 22 half lives and thus we might expect that very few would reach the ground - only about 1 in three million. However, as seen by us, a clock traveling with the muon at 0.98c will appear to be running slow by a factor of 5 and so the effective half life of the muon will be 7.8 microseconds. This is only 4.3 half lives and, as a result about 150,000 out of 3 million will reach the ground and so we can detect a significant muon flux at sea level.

It is worth looking at this as if we were traveling with the muon. We would not see time dilated but the end result *must* be the same. This is achieved because, as seen by the muon traveling at 0.98c, the distance it has to travel is *less* by just the same factor as the time appeared to be dilated to an observer on the ground. This is called **length contraction**. The net result is that the muon still travels for only 4.3 half lives.

Gravitational Time Dilation

If you were an astronaut traveling at a constant speed in a spaceship you would feel weightless, but suppose it accelerates upwards in the direction vertically above you. As the body of the spaceship moved upwards, you would find that your feet would very soon touch and stand upright on the floor and become aware that your body had weight. If the acceleration of the spacecraft was the same as the value of g (the acceleration due to gravity) at the surface of the Earth your apparent weight would be *exactly* the same and you could not tell the difference. Einstein pointed out that there is no way of distinguishing between the two scenarios. The acceleration due to gravity that we experience due to the mass of an object like the Earth is *exactly* equivalent in its effects to those experienced by those within an accelerating frame of reference. This observation became the basis of his General Theory of Relativity.

This concept will enable us to see that there is a second form of time dilation. Imagine our photon clock with the mirrors on each side of the spaceship which is accelerating upwards. If an observer in free space (away from any mass) could see what was going on he would see the photon move horizontally in a straight line and hit the mirror at a point nearer the bottom of the space craft (as the mirror had moved upwards whilst the photon crossed the spaceship). The reflected photon would then cross the spaceship again and he could note its arrival time - the first tick of the clock. The length of this tick as measured by this observer would be exactly the same as if the spaceship were stationary so, as seen by him, the clock would keep the same time as one that was stationary.



The path of a photon crossing back and forth across an accelerating spacecraft as seen by an astronaut in the spacecraft. The path length is longer than if the craft was stationary - the photon clock runs slower.

Now consider what a lady astronaut would observe. She would see that the photon has hit the second mirror at a point nearer to the bottom of the spacecraft but, to her, it will have appeared to follow a (longer) curved path from one side to the other. As Einstein states that this is exactly equivalent to being in a gravitational field caused by adjacent mass, we should expect that, in the presence of matter, light will follow curved lines through space, not straight ones! If she follows the photon back to the mirror on the other side, it will also appear to follow a curved path so that the length of the tick as measured by her will be longer than that observed by our external observer - time has been dilated. This form of time dilation is called Gravitational Time Dilation.

We can carry out a simple calculation to estimate the effect of the curvature. Suppose the photon travels a distance l across the cabin taking a time t = l/c. If the spaceship moving with an acceleration g it will have moved vertically a distance $L = 1/2\ g\ t^2$. (This is a standard formula in simple mechanics.) The angle that the photon appears to be deflected down is given by:

$$\text{theta} = L/l = 1/2\ g\ l\ /c^2 \quad \text{(where theta is in radians)}$$

As one might expect, the angle through which the light has been deflected becomes greater the longer the path it travels in the gravitational field. If we put in the value for g at the surface of the Earth and the width of a typical lecture theatre, say 10m, we get a value of 5×10^{-16} radians or 10^{-10} arc seconds. This could not be observed! However consider the Sun; here, g at its surface is, at 270 m/s^2, about 28 times bigger than on Earth and the light from a star will travel a considerable distance within the Sun's gravitational field (this obviously get less the further away from the Sun). The exact calculation gives a deflection of 1.75 arc seconds - as has now been shown experimentally to a high degree of precision.

As the gravitational field gets stronger the time dilation gets greater as, for example, when approaching a black hole. At what is called the event horizon of the black hole - from within which not even light can escape - the time dilation observed by an observer in free space becomes infinite and time is effectively frozen!

Relativity and the Global Positioning System

Due to time dilation, the atomic clocks providing the time signals in the GPS satellite constellation and traveling around the globe at a speed of 3.9 km/sec, will lose ~7.2 microseconds per day as measured by clocks on the ground. There is, however, an even greater effect due to the fact that the GPS clocks are in a weaker gravitational field: that at the height of GPS satellites is only one quarter that at the surface of the Earth. This makes them run fast compared to clocks on the ground by 45.9 microseconds per day. Combining the two effects give a net offset of +38.7 microseconds per day. If not corrected, this would give rise to an increasing error in position that would increase by ~ 10 km per day. To account for this, the frequency standards on board the GPS satellites are given a rate offset prior to launch, making them run slightly slow - they are set to 10.22999999543 MHz instead of 10.23 MHz. The fact that we can use GPS receivers to navigate only works if both of Einstein's theories are taken into account!

Spacetime

Soon after Einstein produce his theory a very elegant geometrical representation if its ideas was produced by considering what one observed within a 4 dimensional space time - three dimensions of space and one of time and called Minkowski Spacetime. It is worth trying to understand a little about this to perhaps explain why the odd things that are observed happen.

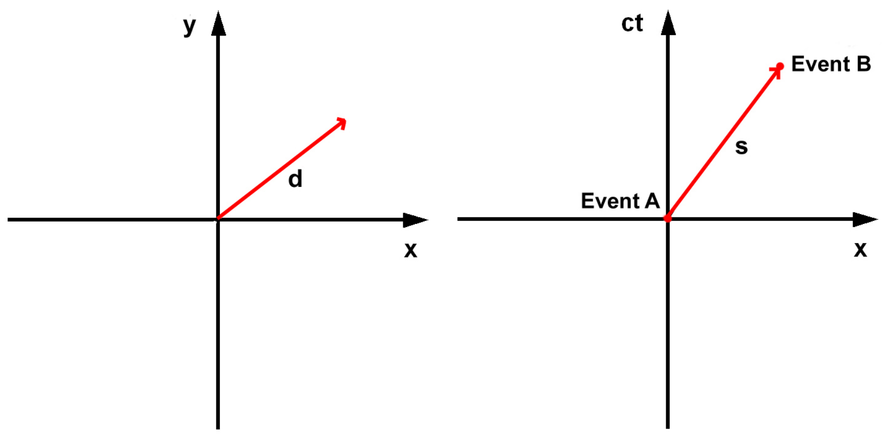
Let us start with a simple analogy: suppose that on a large area of concrete there is a x/y grid marked out (as shown in the accompanying diagram) and a person starts out at the origin and walks for a given time in any direction. Say he will have walked a distance d. At the end of that time his position will lie on a circle of radius d centered on the origin. The route of his path across the concrete will be a straight line having a length (d) and a direction. Things that have both a length (or it can be a speed) *in a specific direction* are called vectors. In physics we use the word "speed" to describe movement in any arbitrary direction i.e., walking at a speed of 3 miles per hour, and "velocity" when a direction is also given i.e., walking due north at a speed of 3 miles per hour. If our walker laid a paint trail behind him, the line would represent the vector of his movement across the concrete. Now one important concept for later is that no matter from where one observed this vector it would have the same length - the length is said to be invariant.

There are two consequent points.

Firstly, one can dissect the vector into two components, one along the x axis and one along the y axis. As the vector length is invariant, if one increases the component along, say, the x axis by making him walk in a different direction then the component along the y direction *must* decrease.

Secondly, the length of the vector is given by Pythagoras's theorem which states that the square of the vector length (d^2) is given by the sum of the x and y components squared so that:

$$\mathbf{d^2 = x^2 + y^2 .}$$



Left: A space only plot. Right: A space-time plot.

How could we translate these ideas into a 4 dimensional space time?

In the case described above we could say that the person leaves the origin at time T = 0 and arrives at a later time, say T = t. So the beginning and end of his walk are two *events* which are determined by both a position in space and a moment in time.

If we can get the geometry right, there *must* be some vector - which we could call the spacetime vector (it is actually called the **Minkoski 4 vector** as it lies within 4 dimensions) - that exists which would also be invariant. This means that, as before, it would have the same direction and magnitude no matter who observed it. A spacetime vector links two "events" in space and time, such as "A" and "B" in the diagram.

The first problem is that we cannot mix different "dimensions" such as length and time. This is easily got round by either converting length into time by dividing by a speed or by converting time into length by multiplying time by a speed. It is simpler to choose the latter so, instead of time, we multiply the time by a constant speed, let's call it c, so the resultant has the dimensions of length. This should not seem too alien as we use "light years" all the time as a unit of length, this being a length given by multiplying the speed of light by the number of seconds in a year.

To make this simpler (without changing the basic idea at all) let's reduce the number of space dimensions to one so we have just one dimension in space (say x) along with one in time multiplied by c, ct.

It turns out that there are only two possible ways of combining these to values to give the magnitude, s, of the spacetime vector:

$$s^2 = (ct)^2 + x^2$$

or

$$s^2 = (ct)^2 - x^2.$$

The first is just Pythagoras again, but it turns out that if this is used to define the length of s, it turns out that some observers would actually see the person arrive before he had left. In fact, in itself, this is not a fundamental flaw but there is a further problem: suppose that this observer saw that when our walker arrived (now only walking in the x dimension) he fell into a pot hole and broke his leg, the observer could actually inform the person *before he left* not to go as far in that direction. This violates the fundamental law of causality - we cannot go back in time to change a future event. The classic example of a problem involving causality is the "grandfather paradox": what if one were to go back in time and kill one's own grandfather before one's father was conceived?

This then, only leaves the second option. However, this still leaves open the possibility of breaching the requirement of causality unless there is a maximum limit to the value at which one can travel through space - a cosmic speed limit if you like.

Let's see what we can learn using the "spacetime vector" formula: $s^2 = (ct)^2 - x^2$.

Suppose you see a friend off for a one hour journey from a station where you are initially both located. We will assume that the railway track is straight so we can work in one dimension and that, at the station, the distance coordinate have the value

0. For you x = 0 and for your friend X = 0. He travels for a time t_1 at a constant speed v. As observed by your friend his position in space will not have changed; X will still be 0, and the time interval that he measured on his wristwatch will be t_1. So, as observed by him:

$$s^2 = (c\, t_1)^2$$

As measured by you, he will have traveled a distance given by vt_2 in a time t_2. Your measurement of the spacetime vector is thus:

$$s^2 = (ct_2)^2 - (vt_2)^2$$

As the spacetime vector is invariant, these must be equal so that:

$$(c\, t_1)^2 = (ct_2)^2 - (vt_2)^2$$

$$c^2\, t_1^2 = c^2 t_2^2 - v^2 t_2^2 \quad \text{or} \quad t_2^2\,(c^2 - v^2) = c^2\, t_1^2$$

Dividing through by c^2 gives:

$$t_2^2(1 - v^2/c^2) = t_1^2$$

So finally we get: $t_2 = t_1 / \mathrm{sqrt}(1 - v^2/c^2)$

or: $\mathbf{t_2/ t_1 = 1 / sqrt(1 - v^2/c^2)}$ **.**

This is exactly the formula we derived earlier but *only* if we interpret c as the velocity of light. So it appears that light travels at the cosmic speed limit.

So, though the spacetime vector linking two events is invariant, as seen by different observers, the time and distance components well normally be different. Our muon experiment described above gives a simple example. As measured from the ground the two events A, when a muon is created in the upper atmosphere, and B when it reaches the ground have a time interval between them of 34 microseconds and a distance interval of 10km but these are 6.8 microseconds and 2 km as observed by the muon.

The Twin Paradox

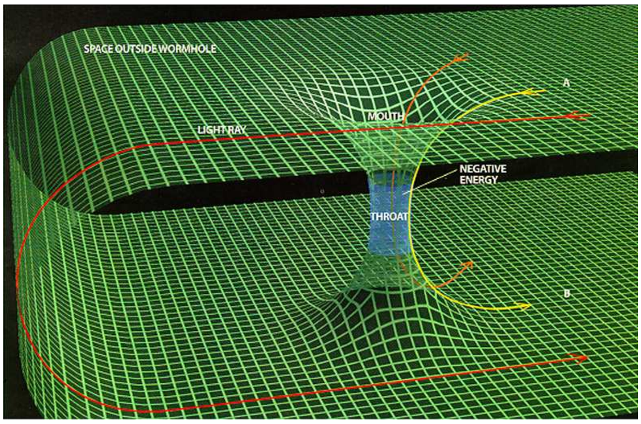
As I am a twin, it might be worth briefly discussing the twin paradox. I stay here on Earth and my twin brother travels into space accelerating into space at 1g - it will feel just like on Earth! After 10 years the ship decelerates at 1 g to come to rest after 20 years. It then turns around and returns home arriving after 40 years as measured by the clock on board the spacecraft. My twin will find that 59,000 years will have passed on Earth and that he has effectively travelled into the future!

The paradox is that the twin in the spaceship could regard himself as stationary and that the earthbound twin as moving away and then back towards him so, by symmetry, they ought to age by the same amount. But of course there is no real symmetry as the space bound twin has accelerated, decelerated and changed direction whilst the earthbound twin has stayed stationary (approximately) on Earth. Some authors state that it is the fact that accelerations take place that breaks the symmetry (i.e., the effects of General Relativity), others that it can be explained simply by the fact that the space bound twin has changed his direction, but all agree that there is no paradox!

Atomic clocks have been flown around the world and are found to have counted fewer "ticks" than a clock stationary on the ground. In 1971, Hafele and Keating sent caesium beam atomic clocks carried by aircraft round the equator in opposite directions. Due to the rotation of the earth, one aircraft travelled round the world at a net speed of about 1500mph, and the other at a net speed of about 500mph. Taking the effects of gravitational time dilation into account the results agreed well with predictions.

Time Travel

As we have shown, special relativity allows one way travel into the future, but is travel into the past possible? As the Minkowski formulation of special relativity is based on the premise that causality must not be violated, then it appears that special relativity bans time travel into the past. But could Einstein's Theory of General Relativity allow a path back in time? This is highly debatable! However, general relativity may possibly allow a loophole if space can be sufficiently "warped" by the presence of matter to form what are called **wormholes** which could, in principle (but probably not in practice) allow a "shortcut" from one part of the universe to another as indicated in the diagram. These are also called "Einstein-Rosen bridges" as in the 1930's Einstein had anticipated them in work he carried out with Nathen Rosen. How might a wormhole (if we could create one) be used to go back in time?



A wormhole producing a shortcut through space

I decide to go to the Andromeda galaxy in my Mk10 spaceship which is parked on the lawn outside my lounge. My wife does not like space travel, but would like to see what Andromeda is like. We make a very short wormhole that goes from our lounge into the spaceship. Though this, she can see what is going on as I accelerate away to a speed of 99.9999999999999999999 the speed of light and reach (as measured on my watch) Andromeda in 4 hours! You might think that the wormhole has to stretch - it doesn't. Amazingly, general relativity allows it to remain the same length throughout the voyage - the further away I am the better a shortcut it is! My wife is able to see

Andromeda as the opening of the wormhole is located conveniently besides a porthole in the spaceship.

I turn the spaceship round and head home arriving on my lawn 4 hour later and so just 8 hours since I left. But everything is different, my house can no longer be seen through the portholes of the spacecraft. However, I am not at all surprised as I had learnt about special relativity from George Gamow's book "Mr Tomkins in Wonderland". I know that by travelling so close to the speed of light on my journey I will have travelled just over 5 million years into the future. (Andromeda is ~2.5 million light years away so, as measured on Earth, had I travelled at the speed of light my return journey would have taken ~5 million years.) I take a look around and leave the spaceship door open. But remember; I still have my link through the wormhole to my lounge. It is time for the supper, so I crawl through the wormhole and greet my wife. *In doing so I travelled back in time 5 million years!*

My journey has turned a wormhole - a tunnel through space - into a tunnel through time and it has become a time machine! People who lived ~5,000,000 years into the future at the location of my house (this would probably still be above sea level as my house is at a height of 500 ft) could enter the spaceship, crawl through the wormhole and travel back to the present. It is probably apparent that a significant limitation of such a time machine is that it is only possible to go as far back in time as the initial creation of the worm hole. This means that using such a machine will not allow you to go back to a time before it was created. As such a time machine has yet to be constructed, tourists from the future cannot reach this far back in time - which perhaps explains why we do not come across them!

The making of such a wormhole would require a substance with negative energy - a form of "exotic matter" - but it appears that quantum physics might make this possible. I would not hold your breath though!

It has been suggested by some physicists that the absence of time travel and the existence of causality might be due to the anthropic principle. The argument is that, if time travel on short time scales is possible, intelligent life could not evolve because it would be impossible for a being to sort events into a past and a future and hence comprehend the world around them.

One final point: if time travel were to be useful, it would have to be a combination of *both* time *and* space travel. If one simply moved forwards in time but did not move in space, then you might find you end up in empty space as the Earth will have moved on in its orbit around the Sun, the Sun will have moved on in its orbit around the centre of the galaxy - which is itself moving through the universe!

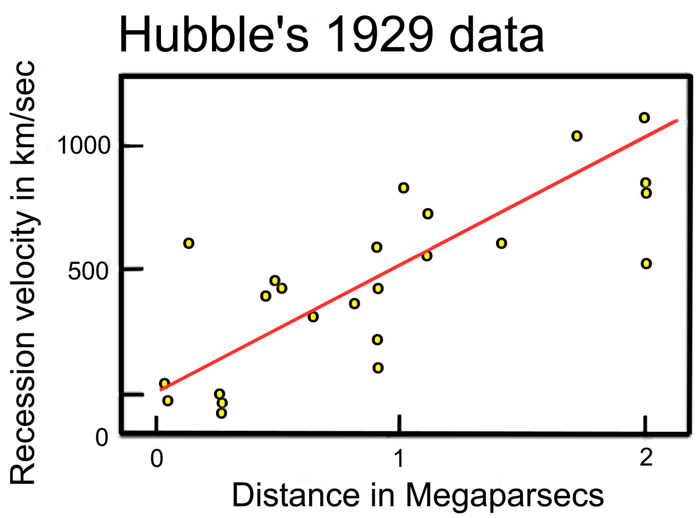
When did time begin?

When the spectra of galaxies were first observed in the early 1900's it was found that their observed spectral lines, such as those of hydrogen and calcium, were shifted from the positions of the lines when observed in the laboratory. In the closest galaxies the lines were shifted toward the blue end of the spectrum, but for galaxies beyond our local group, the lines were shifted towards the red. This effect is called a **redshift** or **blueshift** and the simple explanation attributes this effect to the speed of approach or recession of the galaxy, similar to the falling pitch of a receding train whistle, which we know of as the Doppler effect.

Some of the earliest observations of red and blue shifts were made by the American astronomer Vesto Slipher. By 1915 Slipher had measured the shifts for 15 galaxies, 11 of which were redshifted. Two years later, a further 6 redshifts had been measured and it became obvious that only the nearer galaxies (those within our local group) showed blueshifts. From the measured shifts and, using the Doppler formula, he was able to calculate the velocities of approach or recession of these galaxies. These data were used by Edwin Hubble in what was perhaps the greatest observational discovery of the last century, and it is perhaps a little unfair that Slipher has not been given more recognition.

The expansion of the universe

In the late 1920's, Edwin Hubble, using the 100" Hooker Telescope on Mount Wilson, measured the distances of galaxies in which he could observe a type of very bright variable star called Cepheid Variables which vary in brightness with very regular periods. He combined these measurements with those of their speed of approach or recession (provided by Slipher) of their host galaxies (measured from the blue or red shifts in their spectral lines) to produce a plot of speed against distance. All, except the closest galaxies, were receding from us and he found that the greater the distance, the greater the apparent speed of recession. From this he derived "Hubble's Law" in which the speed of recession and distance were directly proportional and related by "Hubble's constant" or H_0. The value that is derived from his original data was ~500 km/sec/Mpc. Such a linear relationship is a direct result of observing a universe that is expanding uniformly, so Hubble had shown that we live within an expanding universe. The use of the word "constant" is perhaps misleading. It would only be a real constant if the universe expanded linearly throughout the whole of its existence. It has not - which is why the subscript is used. H_0 is the *current* value of Hubble's constant!



Hubble's plot of Recession Velocity against Distance.

If one makes the simple assumption that the universe has expanded at a uniform rate throughout its existence, then it is possible to backtrack in time until the universe would have had no size - its origin - and hence estimate the age, known as the **Hubble Age**, of the universe. This is very simply given by $1/H_0$ and, using 500 km/sec/Mpc, one derives an age of about 2000 million years:

$$
\begin{aligned}
1/H_0 &= 1 \text{ Mpc } / 500 \text{ km/sec} \\
&= 3.26 \text{ million light years} / 500 \text{ km/sec} \\
&= 3.26 \times 10^6 \times 365 \times 24 \times 3600 \times 3 \times 10^5 \text{ sec} / 500 \\
&= 3.26 \times 10^6 \times 3 \times 10^5 \text{ years} / 500 \\
&= 1.96 \times 10^9 \text{ years} \\
&= \sim 2 \text{ Billion years}
\end{aligned}
$$

A problem with age

This result obviously became a problem as the age of the solar system was determined (~ 4,500 million years) and calculations relating to the evolution of stars made by Hoyle and others indicated that some stars must be much older than that, ~ 10 to 12 thousand million years old. During the blackouts of World War II, Walter Baade, recalculated the distance scale and this reduced Hubble's constant to ~250 km/sec/Mpc. There still remained many problems in estimating distances. Gradually the observational data have been refined and, as a result, the estimate of Hubble's constant has reduced in value to about 72 km/sec/Mpc.

If this value is used to calculate the age of the universe we get 13.6 billion years. This is almost exactly the best current value of the age of the universe which is thought to lie between 13.6 and 13.7 billion years. To be honest, this is a lucky coincidence. The universe has *not* expanded at a uniform rate - which our calculation depended on. We now believe that during the first ~9 billion years its expansion rate was slowing - gravity was reigning in the initial expansion - but that for the last 5 billion years the rate of expansion has been increasing. These two effects have canceled out so that now, and only now, in the life of the universe, a linear calculation *does* give the right answer!

Was the Big Bang the origin of time? St Augustine stated that God created the world *with time* not *in time.* Certainly this is true within our own 4 dimensional universe. But some cosmologists believe that our universe was created by the coming together of two “branes” moving in a higher unseen dimension and, if so, time existed before the Big Bang and the cosmos (that is the totality of everything) could be far, far older. We may never know!

Some useful books:

The New World of Mr. Tomkins by George Gamow and Russell Stannard, Cambridge University Press.

Splitting the Second by Tony Jones, Institute of Physics.

The Fabric of the Cosmos by Brian Green, Penguin

Why does $E = Mc^2$ by Brain Cox and Jeff Forshaw to be published in June by De Capo