Vol 32 No. 2 Apr - Jun 2016 ISSN 0970-5953

PHYSICS EDUCATION

Seeds of Black Holes Could be Revealed by Gravitational Waves

www.physedu.in

Volume 32, Number 2

In this Issue

EDITORIAL

This issue comes to you with one more detailed article on gravitational waves even as one more wave was registered by the LIGO detectors. Beyond this, we have other articles ranging from phase transition to special relativity, including one on the height of mountains.

I hope you will enjoy reading this issue. My seasons greetings to all the readers of 'Physics Education'.

M. S. Santhanam
Chief Editor
Physics Education

Einstein's centennial gift: Gravitational waves discovered

Sanjeev Dhurandhar and Sanjit Mitra

Inter University Centre for Astronomy & Astrophysics,
Ganeshkhind,
Pune 411 007, India.
sanjeev@iucaa.in, sanjit@iucaa.in

(Submitted 30-04-2016)

Abstract

A spectacular prediction of Einstein's general theory of relativity is gravitational waves. A century ago - in 1916 - Einstein predicted the existence of gravitational waves. Gravitational waves have now been detected by the LIGO detectors in the US. The physical existence of the waves was established long before by the observations of the Hulse-Taylor binary pulsar whose orbit decays exactly as predicted by Einstein's general theory of relativity. Weakness of the gravitational force implies that the waves are extremely difficult to detect - one must effectively measure distances much smaller than the size of a proton. During the past half century, technology has taken immense strides and the current advanced detectors are now capable of reaching the requisite sensitivity to detect the waves. Gravitational waves carry information about their dramatic origins and about the nature of gravity that cannot be otherwise obtained. A new astronomical window to the universe has been opened. This article will describe the physics of gravitational waves, the technological feats necessary for the detector to achieve unprecedented sensitivities, the current and future global efforts in this direction, the gravitational wave event that was detected, the Indian initiative and contribution to the global effort and the astrophysics that we can learn from this.

1 Introduction

History was created on 11th February 2016, when it was announced that gravitational waves (GW) had been directly detected [1]. The two LIGO detectors of the US detected gravitational waves emitted by the collision of two black holes of about 30 solar mass each on 14th of September, 2015 at 9:50:45 UTC which is at about 15:20 hrs Indian standard time. The data from both detectors, one in Louisiana and the other in Washington state, clearly shows almost identical waveforms in both detectors with time difference of about 7 milliseconds which is consistent with the geographical separation of 3000 km (10 ms GW travel time) between the detectors. This infact marks a three fold discovery: (i) direct detection of GW, (ii) direct detection of black holes and (iii) detection of a black hole binary system. The impact of the discovery is enormous on astronomy and generally on science. It has not only detected black holes but has confirmed general relativity in the strong field regime. It has given rise to the birth of a new astronomy - *Gravitational Wave Astronomy* by opening a new window to the universe. The GW window will complement other windows, namely, the optical one opened by Galileo four centuries ago, radio, infrared, ultraviolet, X-ray and γ-ray in electromagnetic astronomy and also the neutrino. Whenever a new window has been opened it has brought with it unexpected discoveries. Thus it is not unreasonable to expect the unexpected and all the more, because now even the physical interaction is a different one, namely, that of gravity.

The key to gravitational wave detection is the very precise measurement of small changes in distance. In the 1960s, Joseph Weber began his efforts to detect gravitational waves [2]. In a decade of pioneering experiments he investigated resonant bar detectors which were suspended, seismically isolated, aluminium cylinders. His work, though inconclusive encouraged others to build next generation detectors, namely, cryogenic resonant bars and laser interferometric detectors of arm lengths of tens of metres. However, it was soon realised that there were inherent limitations to the design of bar detectors in terms of scalability and narrow band response. A better design was a laser interferometric design which was naturally suited to the quadrupolar nature of GW waves. For laser interferometers, the precise measurement is the distance between pairs of mirrors hanging at either end of two long, mutually perpendicular vacuum chambers. A GW passing through the instrument will shorten one arm while lengthening the other. By using an interferometric design, the relative change in length of the two arms can be measured, thus signalling the passage of a GW. However, the distance measurements are phenomenally small - one-thousand'th the size of a proton! And performing such incredibly small measurement is in fact a feat in technology requiring long arm lengths, high laser power, and extremely well-controlled laser stability [3].

In this article we will first describe the physics of GW, laser interferometric detectors and noise sources, the recent detection of the GW event, the astrophysics we expect from

the new astronomy and the future global network of detectors which includes the detector in India - LIGO-India [4].

2 From Newton to Einstein

The theory of gravitation one usually learns at first is Newton's theory of gravity and the inverse square law. Newton's theory not only explained terrestrial gravity - the legendary falling of an apple - but also the motions of astronomical objects such as the planets and the moon, and in particular Kepler's laws. It came to be known as the universal theory of gravitation because it unified terrestrial gravity with gravity in space as applied to astronomical objects. Its range extended from macromolecules to galaxies and was a resounding success. So then why do we need another theory of gravity?

In 1905, Einstein presented to the world his special theory of relativity [5]. The special theory of relativity essentially deals with measurements of distance, time, mass etc. when the experimenter is moving with respect to the system of objects he is measuring. The principle of relativity says that the laws of physics must be the same for all observers moving uniformly with respect to each other. Special relativity does not concern itself with any specific physical law but requires all physical laws to conform to it. Thus classical mechanics, electromagnetism and quantum physics should obey special relativity. And gravity is no exception. Newton's theory of gravity is not consistent with the special theory of relativity; it is simply unacceptable to have physical theories inconsistent with each other. For example, special relativity requires that all signals must travel at finite speeds, in fact less than or equal to the speed of light in vacuum. But the gravitational force field as described by Newton's theory by the inverse square law is instantaneous, that is, there is no propagation of gravitational forces; the field equations of Newton's theory do not contain *time* - the inverse square law has no time in its description. Thus from this conceptual point of view a new theory of gravity was needed in which gravitational interaction propagates at finite speeds. Although at the beginning of the last century, it was observed that there was a discrepancy in Mercury's orbit - the advance of perihelion of Mercury - Einstein was more concerned with the conceptual problem. Einstein's theory of gravity, the general theory of relativity (GTR), incorporates the special theory of relativity. More importantly, it has come out in flying colours in all gravitation experiments conducted so far - the observations match the theory. Instead of just tinkering with Newton's theory, Einstein formulated conceptually a completely different theory - the general theory of relativity which is also a theory of gravitation [6].

We describe the theory in a prescriptive manner. Matter and energy (described by the energy momentum stress tensor) *curve* the spacetime in its vicinity. Gravitation is the manifestation of the curvature of spacetime. Note that it is a four dimensional curvature - the *spacetime* is curved - and that

space and time have already become a single entity in special relativity. So for instance, if we consider our solar system with the Sun as a central body producing the gravitational field and planets responding to this field and orbiting around it, in Einstein's theory, the Sun *curves* the spacetime around it and the planets move along the straightest possible paths they can in this curved geometry of spacetime. So the orbit of the planet appears curved because the spacetime is curved. The planet strives to follow a "straight" path, but since the spacetime itself is curved and so the "straight" path appears curved. Compare the situation with a sphere. A sphere is an example of the simplest curved space. On the sphere the great circles are the "straightest" possible paths - but they are remarkably different from the straight lines of Euclid's geometry. Such paths are called geodesics.

It remains to prescribe how the matter distribution curves spacetime [7]. This is accomplished by Einstein's field equations[1],

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}\,, \qquad (1)$$

where on the left-hand side (LHS) we have terms describing the curvature in terms of the Riemann tensor and the metric and on the right-hand side (RHS) we have the stress tensor of the matter distribution. The constants G and c denote respectively the Newton's gravitation constant and the speed of light. On the LHS appear the Ricci tensor $R_{\mu\nu}$ and the scalar curvature R, which are derived from the Riemann tensor $R_{\mu\nu\lambda\sigma}$. These equations are the analogue of Newton's equation:

$$\nabla^2\phi = 4\pi G\rho\,, \qquad (2)$$

where ϕ is the Newtonian gravitational potential and ρ is the mass density of matter. But Einstein's equations are much more complicated. They are 10 *coupled* non-linear second order partial differential equations for 10 components of the metric tensor $g_{\mu\nu}$ - the metric tensor is symmetric and so has 10 independent components in 4 dimensions (actually the number of independent equations reduces to six because we have four degrees of freedom in the choice of coordinates). The situation is far more complex than Newton's equation or even Maxwell's equations of electrodynamics and therefore the equations are extremely difficult to solve. Unless one assumes enough symmetries, which effectively reduces their complexity, solutions are hard to come by. For example, no exact analytic solution so far exists for the two body problem in GTR. It is only after years of clever hard work and only recently, that progress has been possible. Solutions can be obtained by a combination of methods involving post-Newtonian approximations [8, 9], numerical relativity [10] and black hole perturbation theory [11].

Further GTR reduces to Newton's theory of gravitation in the limit of weak fields and slow motion as it must, because a new theory must certainly explain phenomena explained by the old theory in its regime of validity; but the new theory may extend be-

[1]Details on Einstein's equations are available in standard textbooks on GR. However, the readers do not need those details to understand this article.

yond the old theory's regime of validity. This happens with GTR. When velocities are not small compared to the speed of light and when the fields are strong, Newton's theory can no longer describe gravitational phenomena accurately and reliably - the spacetime can no longer be considered as a small deviation from the flat (non-curved) spacetime of special relativity - GTR must be used.

The successes and predictions of GTR are spectacular. GTR predicts the expanding universe, black holes and gravitational waves among many other phenomena. In this article we will concern ourselves mainly with GW.

3 The physics of gravitational waves

Einstein's equations admit wave solutions - this is readily seen if we make a weak field approximation [12]. A weak GW is described by a metric perturbation $h_{\mu\nu}$ in general relativity. Typically, for the astrophysical GW sources which are amenable to detection, $h_{\mu\nu} \sim 10^{-22}$. Consider a spacetime which differs slightly from the Minkowski spacetime of special relativity. So the Minkowski metric will be slightly modified. Writing,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}\,, \tag{3}$$

where $\eta_{\mu\nu} = \mathrm{diag}\{1, -1, -1, -1\}$ is the Minkowski metric tensor and where $h_{\mu\nu}$ is a perturbation on this 'Minkowski background'. To the linear order in $h_{\mu\nu}$, it can be easily shown (after a fair amount of algebra) in a certain gauge - transverse and traceless (TT) [13] - that Einstein's field equations reduce to the wave equations:

$$\Box h_{\mu\nu} = \frac{16\pi G}{c^4} T_{\mu\nu}\,, \tag{4}$$

where the $\Box$ is the D'Alembertian operator. It is apparent from this equation that firstly, GTR predicts GW and secondly, GW travel with the speed of light because the velocity c occuring in the $\Box$ operator is the speed of light as seen below:

$$\Box \equiv \frac{\partial^2}{c^2 \partial t^2} - \nabla^2\,. \tag{5}$$

Thus GW are waves in the metric field $g_{\mu\nu}$. Now the curvature or the Riemann tensor is essentially formed by taking the second derivatives of the metric - a very complicated formula; details of which need not concern us here. Thus GW can be described also as waves in the curvature of spacetime. And it is the curvature which can be measured with the help of test masses and thus is a physical field. Thus, one may use either the metric or the curvature to describe gravitational waves.

We can deduce the properties of GW from GTR. We list them below:

- GWs travel with the speed $c \sim 3 \times 10^5$ km/sec.
- GWs are transverse.
- GWs have two polarisations.

It is easy to visualise the two polarisations in the TT gauge; the $h_{\mu\nu}$ can be expressed in terms of just two amplitudes, h_+

and $h_\times$, called the 'plus' and 'cross' polarisations. The two polarisation states are easily understood, if we examine the effect of the waves on test particles. The test particles are just free test masses. A single free mass particle cannot detect a wave (or any gravity) because of the equivalence principle of GTR - one can just transform to the freely falling frame of the test particle and the particle then will remain at rest in this frame and thus will not detect any GW. We need at least two spatially separated particles to observe the effect of GW; one tracks the variation in the separation between the particles as a function of time. Since in GTR the metric is a second rank tensor, it is customary to take a ring of test particles and take the reference particle at the centre. If a weak monochromatic gravitational wave of + polarisation is incident on a ring of test-particles, the ring is deformed into an ellipse as shown at the top of Figure 1. Phases, half a cycle apart, of the GW are shown in the figure. For the × polarisation the ellipses are rotated by an angle of 45°. A general wave is a linear combination of the two polarisations.

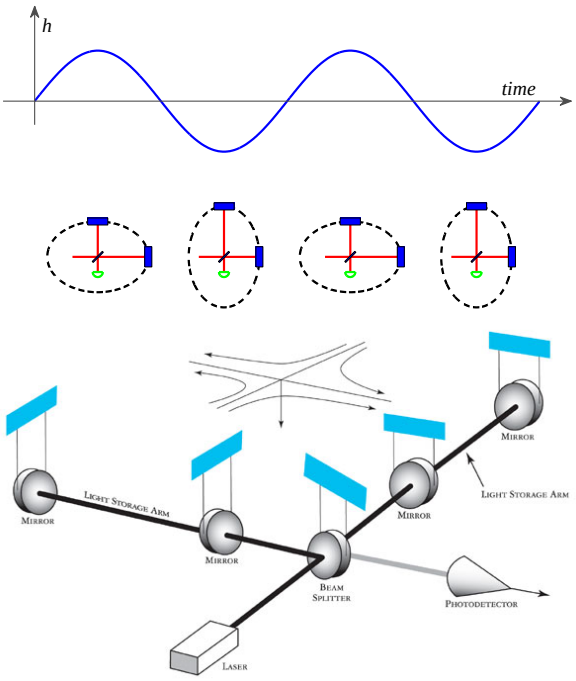


Figure 1: Upper: A circular ring of test particles is deformed into an ellipse by an incident GW (image taken from [14]). Phases, half a cycle apart are shown for the + polarisation. The length change in the interferometric arms is also shown schematically. Lower: a schematic diagram of an interferometer is drawn (Image: Caltech/MIT/LIGO Lab.).

4 Detection of gravitational waves

We will confine ourselves to interferometric detection [3]. The figure 1 shows a schematic diagram of an interferometer. If we select two masses on this ring of test masses at right angles and monitor their distance with respect to the centre of the ring, which we take to be the reference point, we will find that during one half cycle of the wave one arm shortens while the other arm elongates. In the next half cycle of the wave the opposite happens. By using a laser interferometric arrangement a passing GW will produce a time-varying path difference which can be detected on a photodiode.

However, there is a catch! The changes in distances are exceedingly small in astrophysical situations. For example, a neutron star binary at distance of 100 Mpc [2] - a typical distance to a GW source - will produce a differential length change of $\sim 10^{-16}$ cm. for test masses kept few kilometres apart, which is the typical length of the arm of a large scale ground-based interferometric detector! For a GW source, h (a typical component of $h_{\mu\nu}$) can be estimated from the well-known Landau-Lifschitz quadrupole formula. This formula can be obtained by integrating the inhomogeneous wave equation (4) under certain assumptions. The formula relates the GW amplitude h to the second time derivative of the quadrupole moment (which has dimensions of energy) of the source. For obtaining order of magnitude estimates, we can strip the tensor indices of the formula and then it reads:

$$h \sim \frac{4}{r}\,\frac{G}{c^4}\,E^{\rm kinetic}_{\rm nonspherical}, \qquad (6)$$

where r is the distance to the source and $E^{\rm kinetic}_{\rm nonspherical}$ is the kinetic energy in the *nonspherical* motion of the source. If we consider $E^{\rm kinetic}_{\rm nonspherical}/c^2$ of the order of a solar mass and the distance to the source ranging from galactic scale of tens of kpc to cosmological distances of Gpc, then h ranges from 10^{-17} to 10^{-22}. These numbers then set the scale for the sensitivities at which the detectors must operate.

How does the quantity h relate to the change in distance between the test particles? The following formula answers this question. Let L be the distance separating the test masses, then the change in distance δL due to a GW with metric perturbation h is given by,

$$\delta L \sim hL\,. \qquad (7)$$

This result is easily obtained by integrating the geodesic deviation equation. The geodesic deviation equation is justified for ground based detectors because typically the wavelength of the GW - few hundred km or more - is much greater than the distance between the test masses, namely few km, so that the worldlines of the test masses could be thought of as "neighbouring".

However, since detection involves impossibly small measurements, the noise in the detector needs to be suppressed by several orders of magnitude in order that there is a chance of extracting the signal from the noise by statistical signal detection methods. There is a host of noise sources in interferometric detectors which contaminate the data. At low frequencies there is the seismic noise. The seismic isolation is a sequence of stages consisting of springs and pendulums and heavy masses. Each stage has a low resonant frequency about a fraction of a Hz. The seismic isolation acts as a low pass filter, strongly attenuating frequencies much higher than the resonance frequency of the isolation system. This results in a 'noise wall' at low frequencies at around 10 Hz. Also below 10 Hz is the gravity gradient noise which is difficult (if not impossible) to shield. At mid-frequencies upto few hundred Hz, the thermal noise is important and

[2] These are units which astronomers use; 1 Mpc or mega parsec is $\sim 10^{19}$ km.

is due to the thermal excitations both in the test masses - the mirrors - as well as the seismic isolation suspensions. At high frequencies the shot noise from the laser dominates. This noise is due to the quantum nature of light. From photon counting statistics and the uncertainty principle, the phase fluctuations are inversely proportional to the square root of the mean number of photons arriving during a period of the wave. Thus, long arm lengths, high laser power, and extremely well-controlled laser stability are essential to reach the requisite sensitivity.

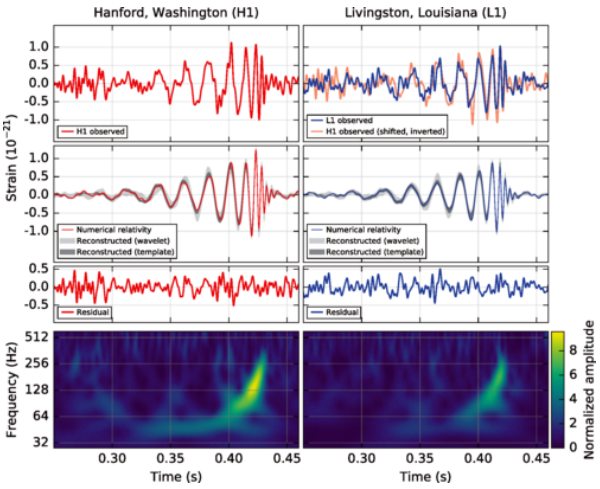


Figure 2: The upper two rows show the GW strains for the two detectors at Louisiana (L1) and Washington State (H1) (image taken from [1]). For visual comparison in the upper right panel H1 data are shown time shifted and inverted to take into account their geographical separation and different orientations. The third row shows residuals and the bottom row shows time-frequency representation of strain data. It is apparent from this plot that the frequency of the signal increases with time.

5 Gravitational waves discovered

The two LIGO detectors of the US in Louisiana and Washington state detected gravitational waves on 14th of September, 2015 at 9:50:45 UTC [1]. The data from both detectors clearly shows almost identical waveforms in both detectors with time difference of about 7 milliseconds which is consistent with the geographical separation of 3000 km (10 ms GW travel time) between the detectors. The waveforms are shown in units of the strain of $h = 10^{-21}$.

The signal was emitted by two black holes of individual masses 28 amd 36 $M_\odot$ which coalesced to form a remnant black hole of mass 62 $M_\odot$ and angular momentum $J = 0.67GM^2/c$, where M is the mass of the final black hole. These masses are given in the source frame. The energy emitted in GW was 3 $M_\odot c^2$ which amounts to $\sim 5\%$ of the total mass. The waveform sweeps through a frequency range from 30 Hz to about 250 Hz in 10 cycles lasting for ~ 0.2 sec. The combined signal-to-noise from two detectors is $\sim$ 24. The estimated distance to the binary black hole is about 410 Mpc. The false alarm probability for the event is less than 2×10^{-7}. The characterstics of the waveform show that this cannot be a neutron star binary nor a neutron

star - black hole binary. Exhaustive investigations were carried out to rule out environmental and instrumental noise. The signal is consistent with GTR.

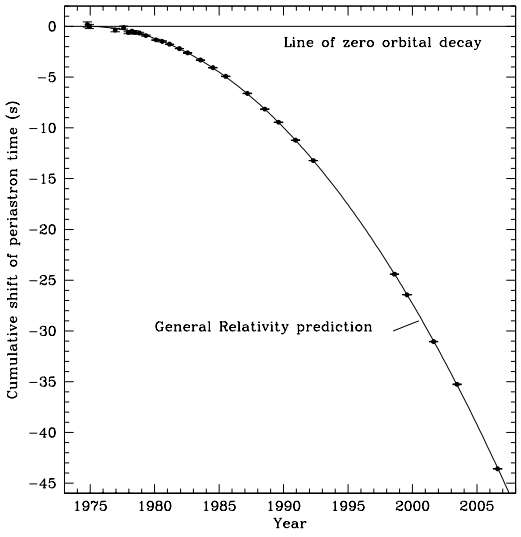


Figure 3: The cumulative advance of periastron is shown as a function of time in years from 1975 to 2005, taken from [15]. The black dots are the observation points while the continuous curve is the prediction of GTR.

Much before this, few decades ago, the existence of GW had been established for which the radio astronomers Hulse and Taylor were awarded the Nobel prize. Hulse and Taylor [16] discovered the binary pulsar PSR 1913 + 16 and subsequent observations showed energy loss and decrease in orbital period of the system [17]. This decay in the orbit is exactly as predicted by GTR. The binary system loses energy through GW, thus shrinking in its orbit and steadily decreasing its period P. The $\dot{P}$ can deduced from the quadrupole formula and also the rate of energy loss through GW. The rate of decrease in the period $\dot{P} \sim -2.4 \times 10^{-12}$ which amounts to about 75 microseconds per year. But because of this the periastron (the epoch at which the stars are closest to each other) advances. The figure above shows the cumulative advance in periastron of the orbit plotted versus the year [15]. The observations exactly agree with the predictions of GTR.

Although this observation establishes the existence of GW it is not a direct detection, because we do not observe the waves themselves; we infer their existence from their effect on the orbit of the binary pulsar.

The direct detection of GW however, has opened a new window to the universe and given rise to the birth of a new astronomy - *Gravitational Wave Astronomy.*

6 Gravitational wave astronomy

6.1 Global network of interferometric detectors

A global network of geographically widely separated detectors is essential for GW astronomy [18] as they are required to (i) localise GW sources in the sky, (ii) increase detection confidence, (iii) increase duty cycle

and (iv) determine polarisation which would give us information on the orientation of the GW source.

The era of advanced detectors has arrived with the state of the art technology which will be capable of observing GW sources and doing GW astronomy. With these future goals in mind, a radical decision has been taken by the LIGO of the US - to build one of the detectors in India in collaboration with India [4]. The reason for this decision is clear - it is to increase the baseline and have a detector far removed from other detectors on Earth, which has several advantages, such as improving the localisation of the GW source, which can then make it feasible to follow up a GW event with electro-magnetic telescopes. There are two LIGO detectors of armlength 4 km in the US, one at Hanford, Washington and one at Livingston, Louisiana geographically separated by 3000 km. The first observation run called O1 of the two LIGO detectors has just taken place which lasted for 4 months. The detectors operated at a sensitivity few times better than the initial detectors. The goal in the next few years will be to improve the sensitivity few times, thereby increasing the volume of the universe the detectors are sensitive to by about an order of magnitude.

In Europe the VIRGO project of Italy and France has constructed a 3 km armlength detector. After commissioning of the project in 2007, it also had science runs. The GEO600 is a German-British project, whose detector has an armlength of 600 metres and is constructed near Hannover. One of the goals of GEO600 is to develop advanced technologies required for the next generation detectors with the goal of achieving better sensitivity.

Japan was the first (around the year 2000) to have a large scale detector of 300 m armlength - the TAMA300 detector under the TAMA project - operating continuously at high sensitivity. Now Japan is constructing a cryogenic interferometric detector called the KAGRA. The purpose of the cryogenics is to reduce the thermal noise in the mirrors and the suspensions and thus increase sensitivity at midrange frequencies.

6.2 The IndIGO consortium and the LIGO-India project

GW research in India had a 25 year legacy and wide recognition in the international GW community. Two groups at IUCAA, Pune and Raman Research Institute (RRI), Bangalore contributed significantly to the global effort, mainly in data analysis of GW signals buried in the noisy detector data and computation of the inspiraling compact binary waveform employing post-Newtonian methods. A lot of trained manpower was created from the students and postdoctoral fellows from these two groups and currently they are occupying key faculty positions both in India and abroad. Given this background, a consortium called *Indian Initiative in GW observations* (IndIGO) was formed in 2009. The aim of this consortium is to foster and promote GW research in India, interact actively with the international community and build up a community of Indian scientists compe-

tent in GW research in theory and experiment. The consortium has proposed a GW detector - LIGO-India - in collaboration with the US, on Indian soil.

The LIGO-India project has been recently approved in principle. The goal of the project is constructing and then operating an advanced interferometric gravitational wave detector in India. A timely opportunity to leapstart gravitational wave research and astronomy in India has arisen through the possibility of the LIGO Laboratory offering and National Science Federation (NSF), US agreeing to transfer one set of components prepared for the advanced LIGO-interferometer, as part of the collaborative effort. Indian scientists will install and operate the detector as well as build the entire infrastructure including the ultra-high vacuum vessels and tubes required to house the interferometer at a suitable, gravitationally and seismically quiet site in India and operate it as part of the global network of detectors for gravitational wave astronomy during the next two decades. The proposal to build and operate the Indian detector is timed to be in this exciting decade of the first detection and observations of GW. To be a key partner in this global endeavour with an interferometer detector built and operated in India is the goal of this project.

6.3 GW astrophysical sources

Several types of GW sources have been envisaged [19, 20] which could be directly observed by Earth-based detectors: (i) burst sources – such as binary systems consisting of compact objects such as neutron stars and/or black holes in their inspiral, merger and ring down phase; burst sources such as supernovae – whose signals last for a time much shorter, between a few milli-seconds and a few minutes, than the typical observation time; (ii) stochastic backgrounds of radiation, either of primordial or astrophysical origin, and (iii) continuous wave sources – e.g. rapidly rotating non-axisymmetric neutron stars – where a weak sinusoidal signal is continuously emitted.

We will discuss here only the compact coalescing binary sources because this is the type of source which has been detected - a black hole binary. Compact coalescing binaries emit enormous amount of GW energy, and also they are clean systems to model; the inspiral waveform can be computed accurately to several post-Newtonian orders [8, 9] adequate for optimal signal extraction techniques such as matched filtering to be used. In the past decade IUCAA has focussed on the design, validation and implementation of search algorithms for inspiraling binaries [21, 22]. Numerical relativity has been able to make a breakthrough by continuing the inspiral waveform to the merger phase and eventually connect it with the ringdown of the final black hole [10]. The full waveforms are obtained by stitching together the inspiral, merger and ring down waveforms. The full waveform consisting of inspiral, merger and ring down can also be obtained directly from numerical relativity alone, but this is computationally very expensive at the moment.

The astrophysical inferences from the currently detected event are as follows [1]. Stellar mass black holes of more than $25M_{\odot}$ ex-

ist and also form binaries within a Hubble time. From the data and the current event, one may estimate the median rate of such events, which turns out to be about 16 events per Gpc^3 per year for a false alarm rate of one per century. One can also deduce how much stochastic background can be produced from the above event rate. An upper limit of 1.2×10^{-22} eV/c^2 can be put on the mass of the graviton from dispersion arguments.

7 The road ahead

Given the situation that the detectors sensitivity will improve in the next few years, one expects to detect many more sources such as the black hole binary event already detected. From these detections we may be able to learn many new aspects in astrophysics, such as the population of black holes of various masses, their distrbution, the GW stochastic background they may produce etc. Also we should be able to observe neutron star - neutron star and neutron star - black hole binaries. We could also have detections of continuous wave sources such as isolated spinning neutron stars, accreting neutron stars etc.

A new window to the universe has been opened and this may bring to us new type of astrophysical sources never imagined. Whenever a new window has been opened, it has brought with it surprises. To cite an example, let us consider radio astronomy. It brought to us pulsars, the cosmic microwave background, radio jets etc. These were completely new discoveries not seen by optical telescopes. Astronomies in other frequency bands have also brought to us new information not available through other windows. This wealth of information from different channels has seen the rise of multi-messenger astronomy where one studies a given astrophysical source pooling together information from the different windows available.

There are also plans to build GW detectors in space. The advantage here is that one can go to very low frequencies of a fraction of a Hz or even mHz. For groundbased detectors a natural limit occurs on decreasing the lower frequency cut-off below ∼10 Hz, because of the gravity gradient noise which is difficult to eliminate below 10 Hz. Thus, the ground based interferometers will not be sensitive below the limiting frequency of ∼10 Hz. But on the other hand, there exist in the cosmos, interesting astrophysical GW sources which should be emitting GW below this frequency such as the galactic binaries, massive and super-massive black hole binaries. If we wish to observe these sources, we need to go to lower frequencies. The solution is to build an interferometer in space, where such noises will be absent and will allow the detection of GW in the low frequency regime. There are plans to build such detectors, such as the eLISA [23] and DECIGO [24] in future. But this may take 20 years or more.

The ground-based detectors and the space-based detectors complement each other in the observation of GW in an essential way, analogous to the way optical, radio, X-ray, γ-ray observations do for electromagnetic waves. As both these types of detectors begin to operate, a new era of GW astronomy is on the horizon and a radically different view of the

Universe is expected to emerge.

Acknowledgments

SM acknowledges the support of Science and Engineering Research Board (SERB), India.

References

[1] B. P. Abbott et al. *Physical Review Letters*, 116(6):061102, February 2016.

[2] J. Weber. *Phys. Rev.*, 117:306–313, Jan 1960.

[3] B. P. Abbott et al. *Phys. Rev. Lett.*, 116:131103, Mar 2016.

[4] Iyer B. et al. LIGO-India, Proposal of the Consortium for Indian Initiative in Gravitational-wave Observations (IndIGO). Internal working note LIGO-M1100296, Laser Interferometer Gravitational Wave Observatory (LIGO), 2011.

[5] A. Einstein. *Annalen der Physik*, 322(10):891–921, 1905.

[6] A. Einstein. *Annalen der Physik*, 354(7):769–822, 1916.

[7] J.V. Narlikar. *An Introduction to Relativity*. Cambridge University Press, 2010.

[8] L. Blanchet, T. Damour, G. Esposito-Farèse, and B. R. Iyer. *Physical Review Letters*, 93(9):091101, August 2004.

[9] L. Blanchet. Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Reviews in Relativity*, 17, February 2014.

[10] F. Pretorius. *Physical Review Letters*, 95(12):121101, September 2005.

[11] C. V. Vishveshwara. *Nature*, 227:936–938, August 1970.

[12] A. Einstein. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 688-696.*, 1916.

[13] B.F. Schutz. *A First Course in General Relativity*. Series in physics. Cambridge University Press, 1985.

[14] B P Abbott et al. *Reports on Progress in Physics*, 72(7):076901, 2009.

[15] J. M. Weisberg, D. J. Nice, and J. H. Taylor. *Astrophys. J.*, 722:1030–1034, October 2010.

[16] R. A. Hulse and J. H. Taylor. *Astrophys. J. Lett.*, 195:L51–L53, January 1975.

[17] J. H. Taylor and J. M. Weisberg. *Astrophys. J.*, 253:908–920, February 1982.

[18] B. F. Schutz. *Classical and Quantum Gravity*, 28(12):125023, June 2011.

[19] Kip S. Thorne. Gravitational radiation. In *300 Years of Gravitation*. Cambridge University Press, Cambridge, UK, 1987.

[20] B. S. Sathyaprakash and B. F. Schutz. *Living Reviews in Relativity*, 12, March 2009.

[21] B. S. Sathyaprakash and S. V. Dhurandhar. *Physical Review D*, 44:3819–3834, December 1991.

[22] S. V. Dhurandhar and B. S. Sathyaprakash. *Physical Review D*, 49:1707–1722, February 1994.

[23] eLISA Mission. https://www.elisascience.org/whitepaper/.

[24] Seiji Kawamura et al. *Classical and Quantum Gravity*, 28(9):094011, 2011.

Travel near the speed of light: an instructive exercise in special relativity

S. Zimmerman

Department of Mathematics
Evergreen Valley College
San Jose, CA 95135

(*Submitted: 23-04-16*)

Abstract

An acceleration of *g* can bring a space vehicle to nearly the speed of light in less than a year. In this paper we study one possibility for such relativistic travel, as measured by both travelers and earthbound observers. Udergraduates who encounter special relativity toward the end of their introductory, calculus based physics course should find the paper's mathematics accessible, and be stimulated by its extension of the usual topics. The sections can be either studied as presented, or offered as problems for investigation and solution by the students.

1. Introduction

The fact that the mass of an object increases without limit as its speed approaches *c* does not of itself prevent the object from attaining a speed arbitrarily close to *c*. In the following examples we investigate changes in time, mass and distance when traveling at relativistic speeds, acknowledging the technical challenges to actually doing so. We will be particularly interested in comparing earthbound measurements with those of the travelers launched into space. With some guidance, students introduced to special relativity in a calculus based physics course should readily understand this study, intended to challenge and stimulate them [1].

The familiar equation $t = v/g$ reveals—perhaps surprisingly—that a constant acceleration of g (9.80665 m/s^2) takes an object from a speed of 0 to c (299,792,458 m/s) in just 30,570,323 *s* or 353.8231829 days. The number of days it takes with an acceleration of g to reach various fractions of c, with no relativistic considerations, is shown in column 1 of table 1.

Although we might adopt any acceleration for this non-relativistic example, we use g because it is familiar to us both intellectually and experientially. Moreover, it requires a force not unreasonable to expect a future space vehicle to maintain, given that even today we can generate a much greater acceleration for several minutes after launch. Thus we will assume that g is not only the constant acceleration for the non-relativistic version of column 1, but also the *initial* acceleration for the two relativistic versions tabulated in columns 2 and 3.

Note that we are *not* assuming that a constant acceleration is maintained. Rather we make a different assumption, one that seems to place a more sensible demand upon the mechanism of a space vehicle than that of constant acceleration. We assume a constant delivery of force—as measured by earthbound observers—one which launches the vehicle with an acceleration of g but which later produces a different acceleration as mass or time are relativistically modified. This force continues until a prescribed speed is attained, after which the vehicle can coast. Of course a model with constant force is only one possibility, although it is surely more practical than a model with constant acceleration, which could demand impossible outputs as mass increases. Students are encouraged to invent their own models with different assumptions.

2. Measurements from the earth, assuming external propulsion and relativistic mass increase

The achievement of relativistic speeds seems more likely to come about through a technology which utilizes an external source of propulsion rather than one which uses onboard fuel [2,3]. Even a cargo of matter-antimatter—the most promising current prospect—would be prohibitively heavy. The student might like to research the possibility of solar sails, and calculate the dimensions needed for feasible travel [4]. The prospects for utilizing "dark energy" seem more remote at this time, but may eventually materialize. Whatever the mechanism, since our study assumes an external source of propulsion, we will encounter no decrease in mass due to consumption of fuel. (We note that the assumption of a diminishing supply of onboard fuel leads to very different scenarios and equations from those studied in this paper.)

Time: Assume that the space-time coordinate system used by observers on the earth is an inertial frame of reference. That is, the accelerations due to the earth's rotation and all other circular motions are negligible and can be disregarded.

In our model, if m_0 is the rest mass of the vehicle, a force of $m_0 g$ is continuously applied until acceleration is no longer needed. As the speed v increases, the mass increases by a factor of $1/\sqrt{1-v^2/c^2}$, as measured by the earthbound observers [5]. Since force is the derivative of the relativistic momentum, we have:

$$\frac{d(m_0 v/\sqrt{1-v^2/c^2})}{dt} = m_0 g, \qquad (1)$$

and after integration we obtain

$$\frac{v}{\sqrt{1-v^2/c^2}} = gt \qquad \text{or}$$

$$t = \frac{v}{g\sqrt{1-v^2/c^2}}. \qquad (2)$$

If we let $u = v/c$ then, with $0 \le u \le 1$,

$$t = \left(\frac{c}{g}\right)\frac{u}{\sqrt{1-u^2}} \qquad (3)$$

The number of days needed to reach various values of u is shown in column 2 of table 1. Compare these with those in column 1, where $t = v/g$. Since

$\frac{u}{\sqrt{1-u^2}} > u$ for 0 ≤ u ≤ 1, the values in column 2 are always greater than those in column 1, with a ratio approaching ∞ as u approaches 1. Thus the vehicle can never achieve the speed of light, but can get arbitrarily close to it. As its mass increases without limit, it takes longer and longer to increase its speed by the slightest increment.

Let us remind ourselves that equation (3) and the values in column 2 specify how observers on earth measure the vehicle's motion. They are not the measurements of the travelers themselves, which we will consider below.

Distance: While the time to attain any particular speed is understandably greater when mass is subject to relativistic increase, is the same true for the distance traveled? If we call the target speed v_t, then the non-relativistic distance traveled from launch until v_t is attained is the familiar

$$s_1 = \int_0^{v_t/g} gt\,dt = v_t^2/2g, \text{ or, if } u_t = v_t/c,$$

$$s_1 = \frac{c^2 u_t^2}{2g} \qquad (4)$$

The distance traveled with relativistic mass increase is

$$s_2 = \int_0^{v_t} v\,dt = gc\int_0^{t_t} \frac{t}{\sqrt{c^2+(gt)^2}}\,dt\,.$$

The upper limit of the integral, t_t , signifies $\frac{v_t}{g\sqrt{1-v_t^2/c^2}}$, obtained directly from (2), while the equivalence $v = \frac{gtc}{\sqrt{c^2+(gt)^2}}$ is an algebraic consequence of (2). After integration

$$s_2 = \frac{c^2}{g}\left(\frac{c}{\sqrt{c^2-v_t^2}} - 1\right),$$

and with $u_t = v_t/c$ this becomes

$$s_2 = \frac{c^2}{g}\left(\frac{1}{\sqrt{1-u_t^2}} - 1\right). \qquad (5)$$

It is easily shown that $s_2 > s_1$ for all $t > 0$. Inspection of (5) reveals that as $u_t \to 1$, $s_2 \to \infty$. This is to be expected, since for the earthbound observers the vehicle will be traveling at a speed near c for an amount of time that according to (3) increases without limit.

3. Measurements by the travelers

Time: How will the situation change if we take into account time dilation, whose effects the travelers' measurements will reflect?

Let T be the elapsed time since launch according to the travelers' onboard clock. T is a function of t, the time as measured by the observers on earth. This function depends upon the vehicle's entire history of motion, not just upon what is occurring at the moment. By contrast, history will prove irrelevant to the travelers' measurement of instantaneous speed, V. That is, we will see that regardless of the vehicle's history, the speed measured by the observers and the speed measured by the travelers are always equal. This of course is an ambiguous statement, but before we clarify what it means, let us show the simple derivation. Let s represent the distance traveled since launch as measured by the observers, and S the distance as measured by the travelers. For any brief interval, time dilation yields:

$$dT = \sqrt{1 - v^2/c^2}\, dt$$

But since the travelers measure the space outside them to be moving past their window at a velocity of v, length contraction yields [6]:

$dS = \sqrt{1 - v^2/c^2}\, ds$, and thus the instantaneous speed measured by the travelers is

$$V = \frac{dS}{dT} = \frac{\sqrt{1 - v^2/c^2}}{\sqrt{1 - v^2/c^2}} \frac{ds}{dt} = \frac{ds}{dt} = v \qquad (6)$$

Time dilation apparently compensates for length contraction, but again, what exactly does (6) assert? We can interpret it with a simple example. Suppose that both the observers and the travelers, with much more knowledge of the solar system and galaxy than we have, specify event A to be the vehicle's passage past a particular comet in the Oort cloud (perhaps even specifying that the vehicle's vector of motion is perpendicular to a line from the vehicle to the center of the comet.) Then the speed v which the observers measure at event A will equal the speed V which the travelers seeon their speedometer at their own event A. That is, the observers and the travelers will agree on the vehicle's speed at the moment event A occurs. There is no confusion as to whether these two equal speeds are "happening at the same time," which is not only confusing but also meaningless. We *can*, however, say something definite about time: as we will show below, if the travelers measure event A happening at time T, and the observers measure it happening at time t, then $T<t$. That is, the travelers measure its occurrence at an earlier time on their clock than the observers do on theirs.

We will use equation (6) to obtain T as a function of t for the specific motion portrayed in (2), but now as measured by the travelers. Without noticing anything strange, since they are not comparing their measurements to those of the observers, the travelers are subject to two relativistic effects: a dilation of time, and a contraction of the space outside as it flies by their vehicle.

Solving for v in (2) gives $v = \dfrac{tgc}{\sqrt{c^2 + (gt)^2}}$, and substituting this into

$dT = \sqrt{1 - v^2/c^2}\, dt$ yields

$$dT = \sqrt{1 - \frac{t^2 g^2 c^2}{c^2(c^2 + (gt)^2)}}\, dt = \sqrt{\frac{c^2 + t^2 g^2 - t^2 g^2}{c^2 + (gt)^2}}\, dt = \frac{c}{\sqrt{c^2 + (gt)^2}}\, dt .$$

Integrating and solving for the constant of integration gives:

$$T = \frac{c}{g} \sinh^{-1}\left(\frac{gt}{c}\right), \text{ or} \qquad (7)$$

$$t = \frac{c}{g} \sinh\left(\frac{gT}{c}\right) \qquad (8)$$

Replacing t in (2) with t from (8) produces the equation

$\dfrac{c}{g} \sinh\left(\dfrac{gT}{c}\right) = \dfrac{v}{g\sqrt{1 - v^2/c^2}}$, which yields

$$T = \frac{c}{g}\ln\left(\frac{v+c}{\sqrt{c^2 - v^2}}\right) \qquad (9)$$

Letting $u = v/c$,

$$T = \frac{c}{g}\ln\left(\frac{u+1}{\sqrt{1-u^2}}\right). \qquad (10)$$

This is still less than satisfactory, for T on the left side of (10) represents the travelers' measurement of time, while u on the right side is the fraction of c as measured by the observers. However, since for any event A the speed V equals the speed v, u could just as well be equated with V/c, and (10) would be exclusively in terms of the travelers' measurements. The value of T for various speeds u is given in column *3*, table 1. The entire column is in italics to indicate that it represents the measurements of the travelers. $T<t$ for all u, an inequality which grows with increasing u.

A note on measurement: Before calculating the distance as measured by the travelers, let us briefly consider the term *measurement* itself. This is a term used repetitively, not only in this paper but in all physical discussions. While the term conjures up—and indeed almost always signifies—an observation using some calibrated instrument or apparatus, it might be useful to explore the concept somewhat. Consider this line of reasoning. Between the above event A and a later event B, such as the encounter with a second comet, there is an interval of time and distance measured differently by the observers and the travelers. (To be clear, here the word "interval" is *not* being used to designate the invariant quantity $[(c\Delta t)^2 - (\Delta x)^2]^{1/2}$.) Let us say that during this interval the traveler (we will assume female), using her onboard clock and her right index finger placed on her left wrist, calculates that during this interval her pulse has held steady at 60 beats per minute. An observer on earth, with electromagnetic access to the traveler's pulse but using his own clock, will calculate that the traveler's pulse began the interval at, say, 40 beats per minute, and after the vehicle's acceleration through the interval, finished at 35 beats per minute. Between events A and B the total number of beats had to be the same for the traveler and the observer, but the onboard clock measured fewer minutes than the clock on earth. To the observer, the traveler's heart seemed to beat increasingly slowly, while the traveler herself felt quite normal.

Any onboard activity would give a similar result. For example, the observer would measure a male traveler's facial hair to be growing slowly, although by the onboard clock the traveler himself would shave according to his usual schedule. To the observer the traveler's speech would get slower and slower as the vehicle accelerated, as would the appearance of emotionson his face, or even—if the capacity to register this had been discovered—the succession of his thoughts. None of these activities would appear or feel unusual to the traveler. To him the onboard clock itself would advance normally, although the observer would perceive it as slowing down compared with his own clock.

The statement that *time* itself slows down can be misleading, although it does give a sense that something more substantial than a mere appearance is actually happening. One must remember that in this study it ultimately refers to a measurement by a clock in an inertial frame, of activities that are accelerating relative to that inertial frame. (A more complete understanding is beyond the scope of this paper, but a comprehensive and comprehensible reference is [7].)

Distance: Taking a closer look at columns 2 and *3*, is there some contradiction here? If the travelers take only 936.45 days to reach $u = .99$ (column *3*), and the observers measure it as 2483.10 days (column 2), how could the speeds always be the same at every event along the route, such as event *A*? Wouldn't the

travelers measure an average speed much faster than that measured by the observers? We can resolve this difficulty by considering the contraction of exterior distance as measured by the travelers.

How far do the travelers measure that they have gone since the moment of launch until the moment they attain a speed of v_t? Call this distance s_3.

$s_3 = \int_0^t V\,dT$. Using (9), integrating, and replacing v_t/c by u_t yields

$$s_3 = \left(\frac{c^2}{2g}\right)\ln\left(\frac{1}{1-u_t^2}\right). \qquad (11)$$

Substitution of various values of u into (5) and (11) indicates that for $0 < u < 1$, the travelers' distance will be shorter than the distance observed from earth. However, we can be more convincing than this. Consider two specific events, B followed closely by C. The observers and travelers see the same speed at B and the same speed at C, and thus the same difference du_t. The observers measure the distance between these events as ds_2 while the travelers measure it as ds_3. Consider the ratio

$$\frac{ds_3}{ds_2} = \frac{ds_3/du_t}{ds_2/du_t} = \left(\frac{d\left(\frac{c^2}{2g}\ln\left(\frac{1}{1-u_t^2}\right)\right)\Big/du_t}{d\left(\frac{c^2}{g}\left(\frac{1}{\sqrt{1-u_t^2}}-1\right)\right)\Big/du_t}\right) = \frac{u_t/\left(1-u_t^2\right)}{u_t/\left(1-u_t^2\right)^{3/2}} = \sqrt{1-u_t^2}\,. \qquad (12)$$

Since $ds_3 < ds_2$ for all such intervals, integrating to find the distance between launch and any later event will always result in $s_3 < s_2$. Comparing (11) with (5), we find that at u = .5, s_3/s_2 = .93 and at u = .99, s_3/s_2 = .32. In fact, by applying L'Hôpital's rule to the ratio s_3/s_2 , we can show that $s_3/s_2 \to 0$ as $u \to 1$. To sum up, although the time measured is less for the travelers, so is the distance. Thus the equality of speeds shown in (6) is compatible with the briefer travel times of column *3*.

Although we will not consider it here, we might note that once the travelers are moving at a speed near *c* the passage of time is so slow by their measurement that within their lifetime they can reach distant locations in the galaxy or even beyond [8,9].

While the assumption of a constant delivery of force may be too simple and even unrealistic, it generates an interesting question: how does the vehicle's power output—the delivery of energy per unit of time—depend upon time as time itself dilates? That is, how do the travelers measure the delivery of force? We leave this for the motivated student to research.

4. Data

Table 1. Exterior propulsion. Time to reach *u=v/c*, or *V/c* to nearest hundredthof a day. In columns 1 and 2, *t* is time for observers; in column *3*, *T* is time for travelers. c = 299,792,458 m/s, g=9.80665 m/s^2

U=v/c	**1**	**2**	**3**
.9999	353.79	25017.19	1752.03
.99	350.28	2483.10	936.45
.98	346.75	1742.47	812.93
.95	336.13	1076.48	648.13

.90	318.44	730.55	520.91
.80	283.06	471.76	388.71
.70	247.68	346.82	306.87
.60	212.29	265.37	245.25
.50	176.91	204.28	194.36
.40	141.53	154.42	149.90
.30	105.15	111.27	109.52
.20	70.76	72.56	71.73
.10	35.38	35.56	35.50

1:No relativistic effects; $t = v/g$.

2:Relativistic mass increase, no time dilation; as observed from earth; $t = (c/g)\left(u/\sqrt{1-u^2}\right)$.

3: Relativistic mass increase, time dilation; as experienced by the travelers;

$$T = (c/g)\ln\left((u+1)/\sqrt{1-u^2}\right).$$

5. Conclusion: turning student questions into research

Let us look at three examples of questions which open the way to stimulating research. (Like many other instructors, I've finally learned to resist giving too complete an answer.)

A thoughtful student will inevitably reason like this: since the travelers see the earth accelerating away from them, why not reverse the calculations? Now the traveler measures the observer's clock to be going slow. What happens when the vehicle returns to earth 35.28 35.56 35.50and the two clocks are compared side by side? How can they both be slow? Isn't there some contradiction? No, in fact, there isn't. The earth is assumed to carry a fixed, inertial frame of reference, while the vehicle carries an accelerating frame, not only because of its acceleration of *g*, but also because of the further acceleration it requires to turn around and return to earth. Our calculations apply to inertial frames only, and these accelerations would affect the clock on the vehicle but not the clock on earth. The situation is not symmetric. Research the "twin paradox."

Another student might point out the imprecision in our calculations: the effects of general relativity, which we have not included in this study, would surely modify our results. True, but the effects of accelerating at *g* are indistinguishable from those of the earth's gravitational field, which we know from experiment has a negligible effect on mass and time [10]. Our calculations are accurate to several decimal places beyond the hundredths to which we've rounded off.

A third questioner, perhaps a down-to-earth engineering student, might dismiss relativistic travel as mere fantasy, theoretically impossible, beyond any feasible technology. This is the stuff of pop

culture, why bother with it? Perhaps so, but an overly conservative attitude—however well informed—can be too restrictive. At the end of the nineteenth century many scientists believed that physics had reached its culmination. Probably the most celebrated comment to this effect was A.A.Michelson's of 1894 [11]: "The more important fundamental laws and facts of physical science have all been discovered, and these are so firmly established that the possibility of their ever being supplanted in consequence of new discoveries is exceedingly remote." Perhaps every era, satisfied with its accomplishments, believes it has made the ultimate discoveries [12]. This attitude is somewhat tempered today, for recent discoveries—the acceleration of the universe's expansion, or the finding of the Higgs boson in 2012—have clearly posed new fundamental questions. Without apology we can speculate sensibly about the future and make tentative assumptions such as those in this paper.

References:

[1] Serway A, Vuille C 2012 *College Physics9th ed*. (Boston, MA: Brooks/Cole) pp 885-933

[2] http://arxiv.org/abs/1205.2281

[3] http://math.ucr.edu/home/baez/physics/Relativity/SR/rocket.html

[4] Celeda T 2013 A solar sailcraft simulation application *Phys. Ed.***48** #4 pp 472-476

[5] Griffiths D J 2013 *Revolutions in Twentieth Century Physics* (Cambridge: Cambridge University Press) pp 58-60

[6] Feynman R 1997 *Six Not-So-Easy Pieces* (Reading, MA: Addison-Wesley) pp 53-64

[7] Mermin, N D 2005 *It's About Time* (Princeton, NJ: Princeton University Press)

[8] Wheeler J A 1990 *A Journey into Gravity and Spacetime* (New York: ScientificAmerican Library) pp 48–53

[9] Zimmerman S 2010 Contact in an Expanding Universe: an instructive exercise indynamic geometry *Eur. J. Phys*. **31** pp 1368-9

[10] http://www.astronomy.ohio-state.edu/~pogge/Ast162/Unit5/gps.html

[11] Michelson, A A 1903 *Light waves and their uses* (Chicago: University of ChicagoPress) available at http://physics.info/space-time/michelson.shtml

[12] http://www.project-syndicate.org/commentary/the-end-of-physics-

How High Could a Mountain Be?

M.N.Murty

Visiting Faculty, Department of Physics,
Sri Sathya Sai Institute of Higher Learning,
Prasanthi Nilayam – 515134, Andhra Pradesh, India.
E -mail: mnarayanamurty@rediffmail.com

(*Submitted: 03-03-2016)*

Abstract

Mount Everest is the world's highest mountain and it has a height of about 9Km. The estimation of the maximum of height of a mountain on earth using basic physics is done by famous Physicist victor Weisskopf (1908 – 2002). According toWeisskopf, the height of a mountain on earth cannot be more than 44Km. If the height of a mountain is more than this critical value, the mountain would start sinking. During the sinking processes, the directionality of the bonds between the atoms of the solid rock below breaks due to the weight of the mountain. As a result, solid rocks below melt and flow aside. This value was estimated by taking the solid rock below the mountain as SiO_2 and the ratio of liquefaction energy to binding energy of the solid rock was supposed to be same as that of ice. Actually, this maximum value of the height of a mountain is less than 44Km because the rock below is warm and needs less energy to liquefy. Weisskopf has also derived a formula in terms of fundamental constants for calculating the critical height of a mountain on any planet with same material structure as earth, if the number of nucleons in the planet is known.

1.Introduction

Mount Everest is the world's height peak in the eastern end of Himalayas and it has a height of about 9Km. Mountains on earth cannot grow above 9Km. The answer is if a mountain is very high, it would start sinking due to its own weight.

Victor Weisskopf (1908 -2002) was a famous physicist. He had made significant contributions in nuclear physics and quantum electrodynamics. Weisskopf served as professor in physics at MIT from 1946 to 1960. He was the director of CERN from 1960 to 1965. After leaving CERN, Weisskopf returned to MIT in 1965. At MIT he became in1967 head of department of physics, a post he held until 1973.Every summer CERN organizes visits for local high school students. The estimation of the maximum height of a mountain is one of the lectures given to high school students by Weisskopf. According to him, the maximum height of a mountain on earth is around 44Km. Weisskopf emphasized the importance of using basic knowledge in physics as a means to make reasonable estimates.

Weisskopf has also derived a formula in terms of fundamental constants for calculating the critical height of a mountain on any planet with

same material structure as earth, if the number of nucleons in the planet is known.

The rest of the paper is organized as follows. The estimation of the height of a mountain on earth by Weisskopf is given in Section 2. The height of a mountain in terms of fundamental constants is presented in Section 3. Conclusion is given in Section 4.

2. Estimation of the height of a mountain on earth by Weisskopf

The following derivation is from the lecture notes(1969) of Victor Weisskopf.

If a mountain is very high, it would start sinking due to its weight. The force due to the weight of the mountain is sufficient to break the directionality of the bonds between the atoms in the rock. This makes the underlying rock melt and flow aside so that the mountain sinks.

The shape of the mountain is taken as shown in Fig.1. Let the mass of the mountain be M and its height be h. Let it sink by a distance x.

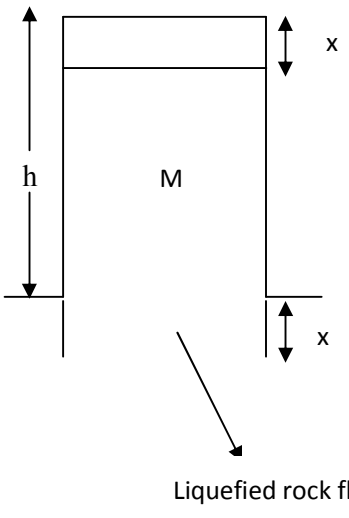


Figure1. Mountain of height h

Loss in gravitational energy when the mountain sinks by x

$= Mgx$ (1)

This loss in gravitational energy is used to liquefy the rock below. If the mountain has to sink by a distance x, a layer of thickness x below the mountain must melt and flow away.

The energy required to liquefy a layer of thickness x

$= E_{liq}(n\,x\,a)$ (2)

Where

E_{liq} = Liquefaction energy (i.e. latent heat of melting) per molecule

$n =$ Number of molecules per unit volume of the rock

$a =$ Area of cross-section of the base of the mountain

Equating (1) and (2), we get

$$Mgx = E_{liq}(n\,x\,a)$$

$$\Rightarrow Mg = E_{liq}\,n\,a \quad (3)$$

Now we conclude that if the mass M of the mountain is less than that given by (3), the mountain will not sink. Hence the masses of stable mountains is given by

$$M \leq \frac{E_{liq}\,n\,a}{g} \quad (4)$$

Let m denote the mass of a molecule of the rock, and A is the number of protons and neutrons in a molecule. If m_p denotes the mass of a proton, then $m \approx A\,m_p$, taking the mass of a neutron approximately equalto the mass of a proton. Then the mass M of the mountain can be written as

M = volume of the mountain $\times$ number of molecules per unit volume $\times$ mass of a molecule

$= ha \times n \times m$

$= h\, a\, n\, A\, m_p$ (5)

Substitution of (5) in (4) gives

$$h \le \frac{E_{liq}}{Ag\, m_p} \quad (6)$$

The above expression shows that the height of a mountain must be less than the critical value $h_c = E_{liq} / Agm_p$ for the mountain not to sink into earth.

When a solid melts, the whole bonds between the atoms are not broken, just the directionality of the bonds are broken. This enables a liquid to flow andthe energy necessary to break the directionality of bonds, i.e to liquefy, is less than the binding energy. To calculate E_{liq}, Weisskopf has compared the latent heat of melting of ice (⊔ 80 cal/gm) with that of latent of vaporisation of water (⊔ 540cal/gm). The latent heat of melting of ice is approximately one-seventh of latent heat of vaporisation of water. Weisskopf has applied this ratio for ice to solids at the baseof the mountain. The binding energy of a solid is the energy applied to a solid to tear it completely into separate atoms. The binding energy of a solid is more than the energy required to tear a liquid into separate atoms. It would be a good estimate to assume that for a solid the melting energy is about one-tenth of binding energy. So Weisskopf takes

Energy of melting ~1/7 energy of vaporisation

~1/10 Binding energy of solid

Hence we have

$$E_{liq} = \beta B \quad (7)$$

where $\beta = \sim 0.1$and $B =$ binding energy of the solid per molecule.

Substituting (7) in (6), we obtain

$$h \le \frac{0.1\, B}{Ag\, m_p} \quad (8)$$

B isabout 2.7ev for SiO_2,the main constituent of the rock. The mass number of silicon atom is 28 and that of oxygen is 16. So, $A = 28 + 2 \times 16 = 60$ for SiO_2.

$m_p = 1.672 \times 10^{-27}$Kg and $g = 9.8\ m/S^2$

Using the above values in (8), we get

$$h \le \frac{0.1 \times 2.7 \times 1.6 \times 10^{-19}}{60 \times 9.8 \times 1.672 \times 10^{-27}} m$$

$\le 43941m$

$\Rightarrow h \le 44\ Km$

Therefore, the height of a mountain on earth must be less than 44Km to be supported by rock at its base. Actually, the upper limit is smaller than 44Km because the rock at the base of the mountain is warm and needs less energy to liquefy. The height of world's top five mountains is given in *Table* 1.

Table1. Five world's highest mountains

Rank	Mountain	Height in Km
1	Everest	8.8498
2	K2(Mount Godwin Austen)	8.6106
3	Kanchenjunga	8.5859
4	Lhotse	8.5161
5	Makalu	8.4630

Mount Everest is the world's highest mountain and it has height of about 9Km. The highest mountain on moon is Mons Huygens, which has a height of 5.5Km. The highest mountain Maxwell Montes on the planet Venus is about 11Km. Boosaule Montes is the highest mountain on Jupiter's moon IO and has a height of 17.5Km. The highest mountain on Mars is Olympus Mons, which has a height of about 22Km. Olympus Mons is about 2.5 times higher than Everest. Hence, the estimation given by Weisskopf for the maximum height of a mountain on earth seems to be valid for other planets as well.

3.Height of a mountain in terms of fundamental constants by Weisskopf

Equation (6) shows that on other planets the maximum height (critical height) of a mountain would be different because the acceleration due to gravity g changes and the planet may be made of different material. So, Weisskopf has expressed (8) in terms of fundamental constants eliminating g .

The force of attraction between the mass m_1 and the earth is given by

$$m_1 g = \frac{G m_1 M_E}{R_E^2}$$

$$\Rightarrow g = \frac{G M_E}{R_E^2} \quad (9)$$

where M_E = mass of the earth

R_E = radius of the earth

G=gravitational constant=6.67×10^{-11}m^3/Kg-S^2

Let N_E = number of nucleons in the earth

Taking the mass of a proton approximately equal to the mass of a neutron, we can write

$M_E=N_E m_p$(10)

The earth consists mostly of SiO_2 (A = 60) and iron (A = 67). The SiO_2 molecule and the iron atom have approximately the same value of A.

Therefore, the number of molecules in earth = N_E/A

The radius of SiO_2 molecule is taken approximatelyequal to that of iron atom. The radius of both SiO_2 and iron is written as fa_o with f□ 4.

Where a_o = Bohr radius = 0.528 Å

The radius of the hydrogen atom in the ground state is called Bohr radius.

Volume of earth = number of molecules in earth × volume of a molecule

$$\Rightarrow \frac{4}{3}\pi R_E^3 = \frac{N_E}{A}\times\frac{4}{3}\pi\left(fa_o\right)^3$$

$$\Rightarrow R_E = \left(\frac{N_E}{A}\right)^{1/3}\left(fa_o\right) \quad (11)$$

Using (10) and (11) in (9), we get

$$g = GN_E m_p\left(\frac{A}{N_E}\right)^{2/3}\left(\frac{1}{fa_o}\right)^2 \quad (12)$$

The *fine structure constant for gravity* is defined as

$$\alpha_G = \frac{G m_p^2}{\hbar c} = 5.9\times10^{-39} \quad (13)$$

This constant is a dimensionless constant.From (13) we get

$$G = \frac{\alpha_G \hbar c}{m_p^2} \quad (14)$$

Where, $\hbar = \frac{h}{2\pi}$

h = Planck's constant =6.62×10^{-34}J-S

c = speed of light in vacuum

Using (14) in (12), we obtain

$$g=\left(\frac{\alpha_G \hbar c}{m_p}\right)A^{2/3}N_E^{1/3}\left(\frac{1}{fa_o}\right)^2 \quad (15)$$

Substituting (15) in (8), the critical height h_c of the mountain is given by

$$h_c=\frac{0.1\times B}{\hbar c}f^2\left(\frac{1}{\alpha_G}\right)\left(\frac{1}{N_E}\right)^{1/3}\left(\frac{1}{A}\right)^{5/3}a_o^2 \quad (16)$$

The binding energy B of the rock can be written as

$B=\eta$ Rydberg

1Rydberg (Ry) $=\frac{m_e e^4}{2\hbar^2}=13.6$ ev

Bohr radius = $a_o=\frac{\hbar^2}{m_e e^2}$

The *fine structure constant* α is defined as

$$\alpha=\frac{e^2}{\hbar c}\approx\frac{1}{137}$$

Where,

m_e = mass of an electron

e = charge of an electron

Using the above values in (16), we obtain

$$\frac{h_c}{a_o}=0.1\eta f^2\left(\frac{\alpha}{2\alpha_G}\right)\left(\frac{1}{N_E}\right)^{1/3}\left(\frac{1}{A}\right)^{5/3} \quad (17)$$

Where, $\eta\approx 1/5$ for SiO_2 .

The above expression (17) gives h_c /a_o in terms of dimensionless constants and it can be used to calculate the critical height h_c of a mountain on any other planet, if the number of nucleons in that planet is known and that planet has same material structure as earth. We cannot use (17) for a celestial object with a material structure different from that of earth as E_{liq} in that case would be different.

4.Conclusion

The maximum height of a mountain on earth according to the estimation of Victor Weisskopf using basic physics is about 44Km. Actually, the maximum height of a mountain is less than 44Km because the rock at the base of a mountain is warm and needs less energy to liquefy. On other planets, the maximum height of a mountain would be different because the acceleration due to gravity g changes and the planet may be made of different material. Weisskopf has also derived a formula in terms of fundamental constants for calculating the critical height of a mountain on any planet with same material structure as earth, if the number of nucleons in the planet is known. Mount Everest is the world's highest mountain and it has a height of about 9Km.The highest mountain on Mars is Olympus Mons, which has a height of about 22Km. So, the estimation given by Weisskopf for the maximum height of a mountain on earth seems to be valid for other planets as well.

References

[1] G.Venkataraman, *Why Are Things The Way They Are? ,* (Universities Press,1992).

Ice Phase Transition as a Sample of Finite System Phase transition

M.R.Khoshbin-e-Khoshnazar

Physics Department,
Curriculum Development Center Organization & Educational Planning,
P.O.Box 15855-363,Tehran, Iran
E-mail: *khoshbin@talif.sch.ir*

(*Submitted: 11-04-16*)

Abstract

A phase is a physically homogeneous part of a substance separated from other parts of the system by an interface. Claim is that a transition from one phase to another at a given pressure occurs at a strictly constant temperature, and e.g. for the ice, it remains solid until it temperature reaches its melting point at 0°C. At that point, the ice stops getting warmer and begins to melt without a temperature change. Experiments show this claim is not correct and should be modified. The discussion concerning phase transitions in finite systems is certainly an interesting subject. Thermodynamic phase transitions are only defined for infinite systems. It has been shown this feature is common aspect in nuclear matter phase transitions, as well.

Key Words: *Phase transition, Ice Phase transition, finite size, percolation, textbook*

Phase Transition

I have reviewed the discussion on phase changes in several introductory physics textbooks [1-7] and found that all of them either imply or state outright that a body will not undergo a phase change until the entire body reaches the transition temperature. This is of course not the case; otherwise an entire ice cube would melt at once rather than from the outside in! Only Young and Freedman [2] hint at this by stating "If we add heat slowly, to maintain the system very close to thermal equilibrium, the temperature remains at 0^oC until all the ice is melted." In the discussion of phase transitions, none of the texts explicitly mentions the possibility of a thermal gradient within the body or that the entire body does not typically change phase simultaneously. When I contacted with Jearl Walker (author of *Fundamentals of Physics* and *The Flying Circus of Physics*) about this matter, he wrote me in private letter :"…of course, I agree with you. What I say in class is that if I hold a blowtorch to an ice sheet…,I can get liquid water running down over the remaining subzero ice. However, I do not to get into that much details in the textbook because *we can not do homework problems*…So, I hope that the

real details, such as you have done, are brought out in upper level thermodynamic courses…"

Actually, some years ago it has been shown in an abandoned paper [8] -by an experiment- that the ice has no definite melting point and is melting at a temperature below 0°C and gradually melts by increasing temperature: The ice has been placed in the cylindrical plastic glass and then the system puts in the water of a big tank (isothermal source). Since the temperature of the water in the big tank does not change significantly when the ice melts, it chose as an isothermal source. To avoid any error they placed thermometer's probe on the center of block ice surface and recorded the temperature as a function of time. The ice initial temperature was -4.2 °C, and that temperature gradually increased to 0.0 °C over a total time of about 15 minutes. This is exactly what one would expect. At time = 0, the entire block was at a temperature of -4.2 °C. The temperature at the probe increased only very slowly since it was insulated from the environment on all sides . As heat from the surroundings was gradually conducted through the ice to the probe, the temperature readings slowly increased, asymptotically approaching 0.0 °. In °C approximately 1/3 of the initial ice volume has been melted**.** While thermodynamics tell us:" in the typical transformation ice(-4.2°C) to ice(0°C),the ice *does not* change (melt) until to reach 0°C and then melts in an *isothermal* process in 0°C (see fig.1)."While their experiment showed that in 0°C one-third of the initial ice volume has been melted and the phase change occurs *in the earlier time and temperature* (in that experiment, approximately in -3.6°C) in an *non-isothermal* process. Results has been shown in Fig.2. In other words, the suggested typical phase transition(fig.1) is not correct and first part of it should be replaced by Fig.2. As it explained, it relates to heat flowing per time across area of the ice for *temperature gradient* which known as *thermal conductivity*.

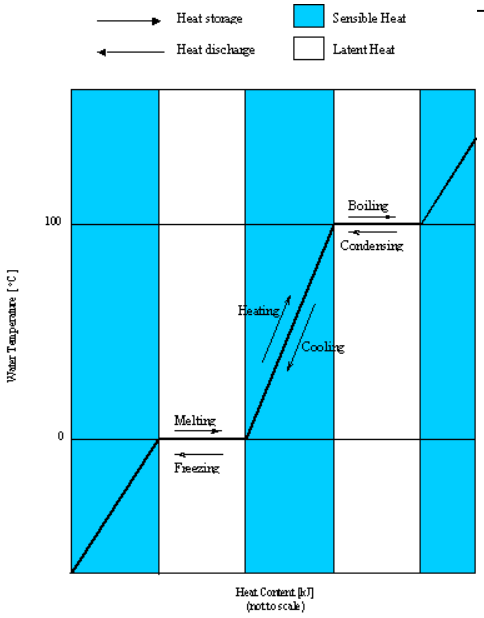


Fig.1 suggested typical phase change for water which introduced in *introductory physics textbooks* . [Adapted from: freespace.virgin.net/m.eckert/new_page6.htm]

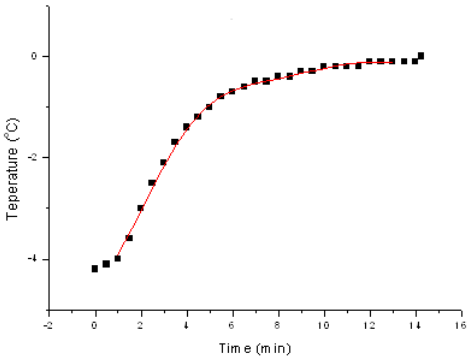


Fig. 2 The temperature vs time data which is fitted by polynomial curve (red curve).

The reason is that the ice at the center would not begin to melt until the temperature there reached 0ºC. In other words, this problem can only be solved by treating differential thickness elements and by recognizing that the ice temperature is not uniform during the phase transition though the temperature at the interface where the melting occurs is fixed at 0^{o}C.

Note that we addressed finite-size and finite-time corrections to the infinite-time bulk phase transition idealization of water. The suggested typical phase change figure is correct in the infinite size limit, given an arbitrary long time scale to establish equilibrium. In other words, the thermodynamic phase transitions are only defined for *infinite* system. Of course, just about everything in a introductory textbook is idealized and full of caveats, not all of which are mentioned and in an introductory treatment it is simply not possible to cover all real-world complications in all physical processes, but the discussion concerning phase transitions in finite systems is certainly an important and interesting subject and it should be noted.

Amazingly, I recently informed about some other papers which suggested same feature is common aspect in nuclear matter phase transitions, as well [9-12]. As we noted already, the strict thermodynamic sense phase transitions are only defined for *infinite* systems- theoretically, systems for which the number of elementary constituents is comparable to Avagadro's number. There are similar investigations of finite-size modifications of nuclear matter phase transitions that they do not just constitute small modifications, but that they dominate the observables. The task of theoretical calculations based on finite discrete lattices is to extrapolate to infinite lattice size. In the nature there are a few systems with numbers of constituents on the order of 10^2 to 10^5.One of them is the fragmentation of atomic nuclei. For decades, there have been speculations that we may be able to see a phase coexistence between the Fermi liquid of ground state nuclei and the hadronic gas phase of individual nucleons and / or small cluster. There is similar feature might also for molecular fragmentation (for Buckyball these studies were undertaken). Another phase transition postulated for finite systems is that between a hadron gas (where quarks and gluons are confined in color singlet) and plasma of quarks and gluons where they can move freely across the entire volume. Under laboratory conditions, this phase transition could be able to explored by central collisions of high-energy heavy ions at beem energies above 10 Gev per nucleon. What all of these mesoscopic systems have in common is that their thermodynamic state variables can not be observed directly. All systems are only transiently excited to the energies that are sufficient to explore the phase transition.

It may be percolation theory [13] could be explained this feature. In such systems, the injection of excitation energy results in constituents moving apart from each other beyond the range of the interaction. This bond breaking causes neighbors to loose contact with each other. Consitituents that are still contacted via bonds will end up as clusters in detectors.

Acknowledgment

I would like to thank Wolfgang Bauer at Michigan state of University, for helpful discussion and a number of suggestions and Jearl Walker at Cleveland State of University for our private communication.

References

1.David Halliday, Robert Resnick and Jearl Walker, Fundamentals of Physics, John Wiley & Sons, 10th editon, 2012

.

2.Hugh D.Young and Roger A.Freedman, University Physics, Addison Wesely Publishing co. 14th edition, 2016.

3.Wolfgang Bauer and Gray D.Wsetfall, University Physics,McGraw-Hill, first edition, 2011.

4.John Cutnell and Keneth Johanson, Introduction to Physics, John Wiley & Sons, 9th edition, 2013.

5.Alan Giambatista and Betty Richardson, Physics, McGraw-Hill,second edition, 2008.

3.Vern J.Ostdiek and Doland J.Board, Inquiry into Physics, Thomson, 5th edithon,2005.
4.Paul A.Tipler, Physics for scientists and engineers, W.H.Freeman, 4th edition,1999.

6.Frank J. Blat, Principles of Physics, Allyn and Bacon, 3th edition,1989.
7.Hanc C.Ohanian, Principles of Physics, W.W.Norton, 3th edition,1994.

8. M.R.Khoshbin-e-Khoshnazar,"Second Look on the Ice Phase Transition". Physics Education (India) **25** ,113(2008).

9. W. Bauer and A. Botvina, "Pre-Equilibrium Particle Emission and Critical Exponent Analysis", Phys. Rev. C **52**, R1760 (1995).

10. W. Bauer, "Extraction of Signals of a Phase Transition from Nuclear Multifragmentation", Phys. Rev. C **38**, 1297 (1988).

11. W. Bauer and S. Pratt, "Size Matters: Origin of Binomial Scaling in Nuclear
Fragmentation Experiments", Phys. Rev. C **59**, 2695 (1999).

12. W. Bauer, "Common Aspects of Phase Transitions of Molecules, Nuclei, and Hadronic Matter", Nucl. Phys. A **681**, 441 (2001).

13. Dietrich Stauffer and Ammon Aharony, *Introduction to Percolation Theory*, Taylor & Francis, second edition, 1994.

Explanation of reflection in Paraxial ray optics by matrix method

Sk. Kalimuddin[(1)] and A.Ghorai[(2)]

Department of Physics; Maulana Azad College
8, Rafi Ahmed Kidwai Road, Kolkata – 700013, India
e-mail :[(1)] premkhan444@rediffmail.com, [(2)] amitavaghorai61@gmail.com

(*Submitted*: 21-03-2016)

Abstract

Matrix method in paraxial ray optics is essential for methodical understanding in graduate physics students. However, there is almost no text book which gives a complete description about the matrix method of paraxial ray optics, in particular, the mechanism of reflection. So in this article the matrix method of reflection from plane and curved surfaces will be dealt with in details and for completeness refraction in optical systems will also be explained.

1. Introduction

Paraxial ray optics is essential for understanding of geometrical optics especially for an optical system consisting of several reflecting and refracting surfaces. Also it is an essential part of undergraduate physics syllabus. However there is almost no standard text book [1–5] which gives a complete description about the matrix method of paraxial ray optics, in particular, the mechanism of reflection although in two recent books on optics [6–7] an introductory outline have been made. So in this article the matrix method of reflection from plane and curved surfaces will be explained in detail and for completeness refraction at plane surface will also be considered.

2. Basic ideas

A combination of reflecting or refracting surfaces or a combination of thin lenses or a thick lens produces an optical system and for lenses it is called lens system. We take x axis as the axis of the optical system and configure this system with small aperture such that only paraxial rays enter the system. These paraxial rays are very close to the principal axis (x axis) and make very small angle with it. For this we take cylindrical symmetry such that a set of paraxial rays lie in a plane making same small angle of inclination.

3. Coordinate of a ray

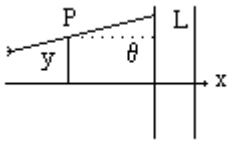


Figure 1

We consider an optical system L and a paraxial ray PL as shown in figure 1. The coordinates of point P can be specified by the distance from the x axis and the inclination of the ray PL with principal axis (x axis) measured in counter clockwise direction (for measurement in the clockwise direction the angle will be negative). The slope of PL is $\tan\theta \to \theta$ and it will vary with medium. So we choose a new coordinate $\lambda = \mu \tan\theta \to \mu\theta$. Thus the coordinates of the paraxial ray at P will be (λ, y), where y is the distance of point P from x axis or principal axis and the refractive index of the medium is μ. Of course, the coordinate at L will

differ inside and outside the optical system because of refractive index.

4. Translation matrix

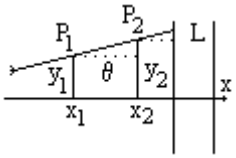


Figure 2

We take a paraxial ray P_1P_2 specified by inclination and distance from principal axis (x axis). The coordinates of P_1 and P_2 will be respectively (λ_1, y_1) and (λ_2, y_2). From geometry shown in figure 2 the slope of paraxial ray P_1P_2 is same and so for $x_2 - x_1 = x$ and for small angles for paraxial rays we get

$$\tan\theta_2 = \tan\theta_1 \Rightarrow \mu\theta_2 = \mu\theta_1 \Rightarrow \lambda_2 = \lambda_1$$

Also

$$y_2 = y_1 + (x_2 - x_1)\tan\theta \rightarrow y_1 + x\theta_1 = y_1 + x\lambda_1/\mu$$

Hence in matrix notation

$$\begin{bmatrix}\lambda_2\\ y_2\end{bmatrix} = \begin{bmatrix}1 & 0\\ x/\mu & 1\end{bmatrix}\begin{bmatrix}\lambda_1\\ y_1\end{bmatrix} = T\begin{bmatrix}\lambda_1\\ y_1\end{bmatrix}$$

Here $\begin{bmatrix}\lambda_2\\ y_2\end{bmatrix}$ and $\begin{bmatrix}\lambda_1\\ y_1\end{bmatrix}$ are two column matrices defined as the coordinate of the paraxial ray P_1P_2 at points P_1 and P_2 respectively. The 2×2 square matrix $T = \begin{bmatrix}1 & 0\\ x/\mu & 1\end{bmatrix}$ is called **translation matrix**. Here x is the translation along x axis of point P_1 to point P_2 and the refractive index of the medium is μ. It should be noted that determinant of translation matrix is unity, i.e. $\det T = |T| = 1$. We can also write the matrix equation as $[2] = T[1]$.

5. Reflection at a plane surface

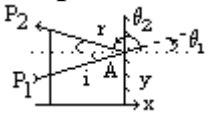


Figure 3

We take an incident paraxial ray P_1A which after reflection at A by a plane mirror produces a reflected ray AP_2 as shown in figure 3. From laws of reflection we know the angle of incidence is equal to the angle of reflection. Also at the point of incidence $P_1 \rightarrow A$ and $P_2 \rightarrow A$. Thus the angles subtended with dotted line parallel to x axis by incident and reflected ray will be $\theta_1 = i$ and $\theta_2 = \pi - r$ (or $\theta_2 = -r$). The negative sign is due to the fact that angle of incidence and angle of reflection are measured in the opposite direction from the x axis or principal axis (angle in counter clockwise direction measured from the dotted line is positive while in clockwise direction is negative). Thus $i = r$ gives

$$\mu\tan i = \mu\tan r \Rightarrow \mu\tan\theta_1 = \mu\tan(\pi - \theta_2)$$

Or, $\lambda_1 = -\lambda_2$

Also $y_2 = y_1 = y$

Hence in matrix notation

$$\begin{bmatrix}\lambda_2\\ y_2\end{bmatrix} = \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\lambda_1\\ y_1\end{bmatrix} = R_F\begin{bmatrix}\lambda_1\\ y_1\end{bmatrix}$$

Here $\begin{bmatrix}\lambda_2\\ y_2\end{bmatrix}$ and $\begin{bmatrix}\lambda_1\\ y_1\end{bmatrix}$ are two column matrices defined as the coordinate of the paraxial ray P_1P_2 at points P_1 and P_2 respectively which are very close to point A on the reflecting plane surface. The 2×2 square matrix $R_F = \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$ is called the **reflection matrix** for plane surface. It should be noted that determinant of reflection matrix is negative unity, i.e. $\det R_F = |R_F| = -1$. We can also write the matrix equation as $[2] = R_F[1]$. This is the essential difference between reflection and refraction.

6.Reflection at a concave surface

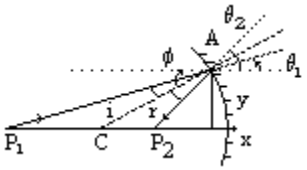


Figure 4

In case of reflection at a concave spherical surface we take an incident paraxial ray P_1A which after reflection at A by a spherical mirror produces a reflected ray AP_2 (figure 4). Also at the point of incidence $P_1 \rightarrow A$ and $P_2 \rightarrow A$. In this case we have to measure angles from principal axis for which a horizontal line is drawn at A. Thus the angles subtended with dotted line parallel to x axis by incident and reflected rays will be $i = \phi - \theta_1$ and $r = \theta_2 - \phi$ respectively.Here ϕ is the angle subtended by AC, the radius of curvature with positive direction of x axis or principal axis. From laws of reflection we know angle of incidence is equal to angle of reflection or

$$\tan i = \tan r \Rightarrow i = r \Rightarrow \mu(\phi - \theta_1) = \mu(\theta_2 - \phi)$$

Or, $\mu\theta_2 = 2\mu\phi - \mu\theta_1$

Or, $\lambda_2 = -\lambda_1 + 2\mu y_1 / R$

Here the radius of curvature of the spherical mirror is R and so $\phi = y_1 / R$.

Also $y_2 = y_1 = y$

Hence in matrix notation

$$\begin{bmatrix} \lambda_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 2\mu/R \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ y_1 \end{bmatrix} = R_F \begin{bmatrix} \lambda_1 \\ y_1 \end{bmatrix}$$

The 2×2 square matrix $R_F = \begin{bmatrix} -1 & 2\mu/R \\ 0 & 1 \end{bmatrix}$ is called **reflection matrix** for concave spherical surface. It should be noted that determinant of reflection matrix is negative unity i.e. $\det R_F = |R_F| = -1$. For $R \rightarrow \infty$ we have plane reflecting surface and $R_F = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. We can also write the matrix equation as $[2] = R_F [1]$.

7. Reflection at a convex surface

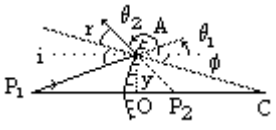


Figure 5

In case of reflection at a convex spherical surface we take an incident paraxial ray P_1A which after reflection at A by the mirror produces a reflected ray AP_2. Also at the point of incidence $P_1 \rightarrow A$ and $P_2 \rightarrow A$. Here we have to measure angles from principal axis for which a horizontal line is drawn at A. Thus from figure 5 the angles subtended with dotted line parallel to x axis by incident and reflected ray will be respectively $i = \theta_1 + (-\phi)$ (ϕ is measured from the dotted line in clockwise direction and so negative) or $i = \theta_1 + (\pi - \phi)$ and $r = (-\theta_2) - (-\phi)$ (θ_2 is negative because it is measured in the clockwise direction)or $r = (\pi - \theta_2) - (\pi - \phi)$. Hence from laws of reflection we know angle of incidence is equal to angle of reflection or

$$\tan i = \tan r \Rightarrow \tan(\theta_1 - \phi) = \tan(-\theta_2 + \phi)$$

[or, $\tan i = \tan r \Rightarrow \tan(\pi + \theta_1 - \phi) = \tan(-\theta_2 + \phi)$]

Or, $\theta_1 - \phi = -\theta_2 + \phi$

Or, $\mu\theta_2 = -\mu\theta_1 + 2\mu\phi \Rightarrow \lambda_2 = -\lambda_1 + 2\mu y_1 / R$

If the radius of curvature of the spherical mirror be R then $\phi = y_1 / R$.

Also $y_2 = y_1 = y$

Hence in matrix notation

$$\begin{bmatrix} \lambda_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 2\mu/R \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ y_1 \end{bmatrix} = R_F \begin{bmatrix} \lambda_1 \\ y_1 \end{bmatrix}$$

The 2×2 square matrix $R_F = \begin{bmatrix} -1 & 2\mu/R \\ 0 & 1 \end{bmatrix}$ is called **reflection matrix** for convex spherical surface. It should be noted that determinant of reflection matrix is negative unity i.e. $\det R_F = |R_F| = -1$. For $R \rightarrow \infty$ we have plane reflecting surface and $R_F = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. We can also write the matrix equation as $[2] = R_F [1]$.

8. Object image relation for spherical mirror

It is now easy to derive the system matrix for an object at a distance u from the surface of the mirror and the corresponding image distance υ in case of reflection at the spherical curved surface as follows:

Object → Translation → Reflection → Translation → Image

$$[2] = T_u[1]$$
$$[3] = R_F[2]$$
$$[4] = T_\upsilon[3]$$

Or, $[4] = T_\upsilon R_F[2] = T_\upsilon R_F T_u[1] = S[1]$

Here $S = \begin{bmatrix} -1+2u/R & 2\mu/R \\ -\upsilon/\mu + 2u\upsilon/\mu R + u/\mu & 2\upsilon/R+1 \end{bmatrix}$

$$\begin{bmatrix} \lambda_4 \\ y_4 \end{bmatrix} = \begin{bmatrix} -1+2u/R & 2\mu/R \\ -\upsilon/\mu + 2u\upsilon/\mu R + u/\mu & 2\upsilon/R+1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ y_1 \end{bmatrix}$$

Thus $\lambda_4 = (-1+\frac{2u}{R})\lambda_1 + \frac{2\mu}{R} y_1$

$$y_4 = (\frac{u}{\mu} + \frac{2u\upsilon}{\mu R} - \frac{\upsilon}{\mu})\lambda_1 + (1+\frac{2\upsilon}{R}) y_1$$

Equating the coefficient of λ_1 to zero (or for axial point object and image $y_4 = y_1 = 0$) we get

$$\frac{1}{\upsilon} - \frac{1}{u} = \frac{2}{r} = \frac{1}{f} = P$$

9. Refraction at a plane surface

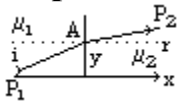


Figure 6

We take an incident paraxial ray P_1A which after refraction at A by aplane interface separating two medium 1 and 2 with refractive indices respectively μ_1 and μ_2, produces a refracted ray AP_2. From laws of refraction we know Snell's law is valid. Also at the point of incidence $P_1 \rightarrow A$ and $P_2 \rightarrow A$. Thus the angles subtended with dotted line parallel to x axis by incident and reflected ray will be $\theta_1 = i$ and $\theta_2 = r$. Thus

$$\mu_1 \sin i = \mu_2 \sin r \Rightarrow \mu_1 i = \mu_2 r \Rightarrow \lambda_2 = \lambda_1$$

Also $y_2 = y_1 = y$

Hence in matrix notation

$$\begin{bmatrix} \lambda_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ y_1 \end{bmatrix}$$

Here $\begin{bmatrix} \lambda_2 \\ y_2 \end{bmatrix}$ and $\begin{bmatrix} \lambda_1 \\ y_1 \end{bmatrix}$ are two column matrices defined as the coordinate of the paraxial ray P_1P_2 at points P_1 and P_2 respectively. The 2×2 square matrix $R_R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is called the **refraction matrix** for plane surface. It should be noted that determinant of refraction matrix is unity, i.e. $\det R_R = |R_R| = 1$. This gives the difference between reflection and refraction as stated earlier. We can also write the matrix equation as $[2] = R_R[1]$.

10. Conclusions

It is now a simple task to derive the refraction matrix for refraction at spherical surface, the object – image relation, the system matrix for lens and lens system, so on. All these are mentioned in the standard text books mentioned above.

Thisformalism clearly depicts the insight concept of matrix method of paraxial ray optics based on the undergraduate theoretical knowledge. Emphasis is given so that undergraduate students can think of new and novel ideas to increase their skill to create new and novel problems for proper understanding.

11. Acknowlegements

The authors gratefully acknowledge the DBT Star College fund for this theoretical project work.

12. References :

[1] Ajoy Ghatak, *Optics,* 4th edition,(Tata McGraw-Hill Pub. New Delhi, 2008).
[2] F. A. Jenkins and H. E. White, *Fundamentals of optics,* 4th ed. (McGraw Hill, Inc, 1976).
[3] R. S. Longhurst, *Geometrical and physical optics*, *(*Orient Longman, Hyderabad, 1973).

[4] Eugene Hecht, *Optics*, 4th ed. Addison Wesley, 2002.
[5] M. Born and E. Wolf, *Principles of Optics*, 5th ed. Pergamon Press, 1975.
[6] F. A. Pedrotti, L. M. Pedrotti and L. S. Pedrottii, *Introduction to optics*, 3rd ed. Pearson, 2014.
[7] S. Prakash, *Optics*, 4th ed. Pragati Prakashan, 2015.

Velocity-surface stability of a moving charged particle in a controlled electromagnetic field

L. Acho[1]

[1]Department of Mathematics
Polytechnic University of Catalunya
Barcelona 08036, Spain.
leonardo.acho@upc.edu

(Submitted 23-12-2015)

Abstract

Given the equation motion of a moving charged particle in a controlled electromagnetic field, this paper proves that its velocity-trajectory motion converges to an specified velocity-surface in the 3-D Euclidean dimensional space. This is basically realized by just manipulating the electric field of an electromagnetic field. Lyapunov theory is invoked to test our statement; besides, a numerical example is provided to support our theoretical contribution. Finally, we consider that the exposition of this paper could be of interest to undergraduate students.

1 Introduction

The second Lyapunov method, also called *The Direct Lyapunov Method*, is now widely used to analyse stability motion of dynamic systems due to its simplicity and efficiency. This theory has been applied, for instance, in physics, astronomy, chemistry, biology, and so on. The main benefit of Lyapunov theory is its ability to conclude stability of systems (*stability in the sense of Lyapunov*) without explicitly integrating the differential equations involved in our system. This method is basically a generalization of the idea that a energy function, the *Lyapunov function*, associated to a system to be analysed, is decreas-

ing along the system trajectory. For a brief discussion on this theory, see, for instance [1].

On the other hand, and motivated by applications where the natural operation mode is periodic, for instance, orbital stabilization of mechanical systems (see, e.g., [2] and references therein), stability analysis of a velocity-surface for a moving charged particle in a electromagnetic environment is here analysed by using Lyapunov theory. That is to say, given the motion of a travelling charged particle in a controlled electromagnetic field, we prove that its velocity-trajectory converges to an specified velocity-surface in the 3-D Euclidean dimensional space. At last to say, the use of a Lyapunov function to conclude orbital stability of dynamical systems is also covered in [3].

The rest of this paper is organized as follows. Section II gives the problem statement and a solution to it. Section III shows a numerical experiment to support our theoretical affirmation. Finally, Section IV states the conclusions.

2 Velocity-surface stability analysis

The objective of this section is to demonstrate that it is possible to produce a 3-D-velocity-surface such that any velocity-trajectory of a charged particle in motion starting close enough to this surface, will converge to it as time goes on. This is realized by manipulating the electric field of an electromagnetic field. Lyapunov theory is the mathematical tool employed to stablish this statement.

The equation describing the motion of a charged particle Q in a electromagnetic fields is defined by

$$m\frac{d\mathbf{v}}{dt} = Q[\mathbf{v}\times\mathbf{B}+\mathbf{E}], \tag{1}$$

where $\mathbf{B}$ and $\mathbf{E}$ are the magnetic and electric fields, respectively; $\mathbf{v}$ is the velocity of the charged particle in this electromagnetic environment, and m is the particle mass. Let us suppose that the electric field (a controlled one) is

$$\mathbf{E} = \mathbf{v}(k- \mid \mathbf{v} \mid), \tag{2}$$

where $|\cdot|$ is the vector Euclidean norm, and $k \in R$ is a positive parameter at our disposal. Then, equation (2) in (1) produces

$$m\frac{d\mathbf{v}}{dt} = Q\left[\mathbf{v}\times\mathbf{B}+\mathbf{v}(k- \mid \mathbf{v} \mid)\right]. \tag{3}$$

Let us use the next candidate Lyapunov function:

$$V(t) = \mathbf{v}\cdot\mathbf{v}. \tag{4}$$

Then, its time-derivative along the system trajectory (3) yields

$$\begin{aligned}\dot{V}(t) &= \frac{dV(t)}{dt}\\ &= \frac{2Q}{m}\left[\mathbf{v}\cdot(\mathbf{v}\times\mathbf{B})+\mathbf{v}\cdot\mathbf{v}(k- \mid \mathbf{v} \mid)\right]\\ &= \frac{2Q}{m}\left[\mathbf{v}\cdot\mathbf{v}(k- \mid \mathbf{v} \mid)\right]. \end{aligned}\tag{5}$$

Observe that $\dot{V}(t)$ is positive if $\mid \mathbf{v} \mid < k$ and negative if $\mid \mathbf{v} \mid > k$. This means that the velocity-surface equation (VSE), $\mid \mathbf{v} \mid = k$, is locally attractive; i.e., for any velocity-trajectory starting close enough to this velocity-surface, the system trajectory (3) will converge to it. On the other hand, if the origin of the system (3) is the only equilibrium point; then, any trajectory, but the origin, will converge to the VSE[1]. In resume, we arrive to the following result.

Theorem 1.- Given the dynamic of a charged particle in motion (1), its velocity-trajectory will locally converge to the velocity-surface established by $\mid v \mid = k$, $k \in R^+$, if we set the electric field as $\mathbf{E} = \mathbf{v}(k - \mid \mathbf{v} \mid)$ in an electromagnetic field. Moreover, if the system only has the origin as its unique equilibrium point; then, any velocity-trajectory starting anywhere but the origin will converge to the stated velocity-surface.

3 Numerical example

This section describes a numerical example. Hence, if $\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k}$, and $\mathbf{B} = B_x\mathbf{i} + B_y\mathbf{j} + B_z\mathbf{k}$, then the corresponding ordinary differential equations of (3), yields

$$\begin{aligned} \dot{v}_x &= \frac{Q}{m}(B_z v_y - v_z B_y + p v_x), \\ \dot{v}_y &= \frac{Q}{m}(B_x v_z - v_x B_z + p v_y), \\ \dot{v}_z &= \frac{Q}{m}(B_y v_x - v_y B_x + p v_z), \end{aligned}$$

where $p = k - \sqrt{v_x^2 + v_y^2 + v_z^2}$. Using $k = 2$ and the charged particle with $Q = 8\mu$c and mass $m = 0.5\mu$ kg, simulation results of the above system are shown in Figures 1, 2 ad 3. Each figure present the same experiments using trajectories generated by three different sets of initial conditions. The *blue* line corresponds to $v_x(0) = -3.5$, $v_y(0) = 0.0$, and $v_z(0) = 3.5$; the *red* line for the case when $v_x(0) = 0.0$, $v_y(0) = 0.0$, and $v_z(0) = -1.5$; and the *yellow* one with $v_x(0) = 0.0$, $v_y(0) = -4.0$, and $v_z(0) = 0.0$. All of them in meter per second. Finally, a free space motion is assumed. From these figures, we can appreciate that these trajectories converge to the expected velocity-spherical-surface (the green surface shown in the mentioned figures).

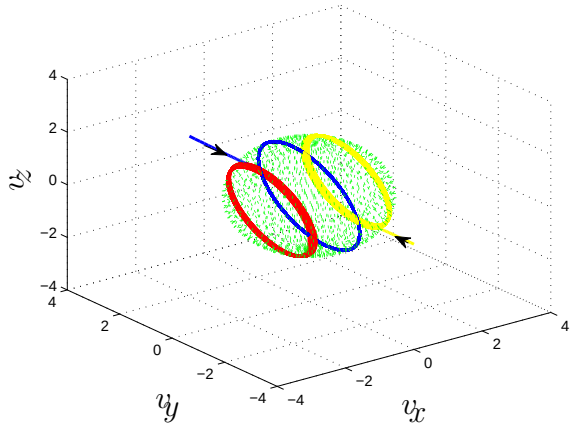


Figure 1: Simulation result using $B_x = B_z = 0.02$T and $B_y = -0.02$T.

[1] Almost the same lines of thinking are employed by [3] to conclude orbital stability in one of the given examples.

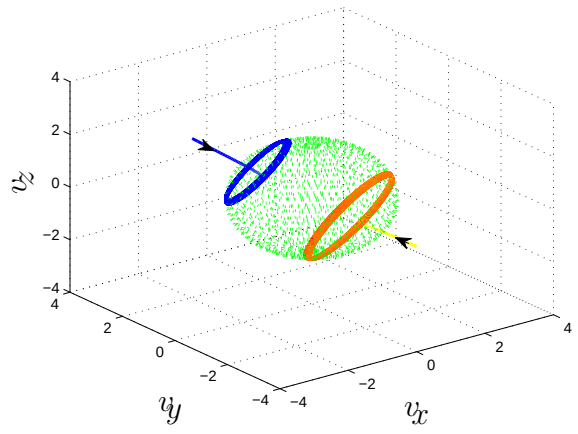


Figure 2: Simulation result using $B_y = B_z = 0.02\text{T}$ and $Bx = -0.02\text{T}$.

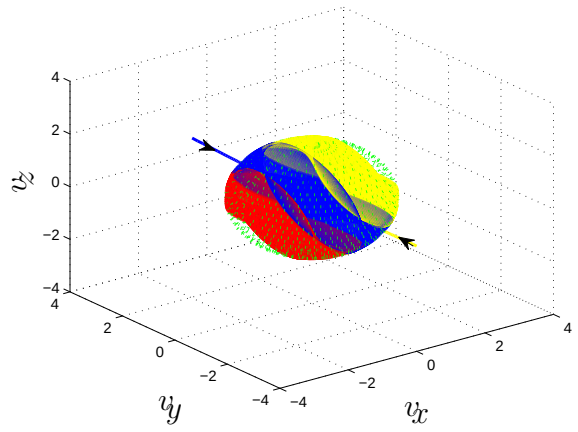


Figure 3: Simulation result using $B_x = B_z = -0.02\text{T}$ and $B_y = 0.02\cos(0.002t)\text{T}$.

4 Conclusion

Lyapunov theory was used to prove stability of a specified velocity-surface of a moving charged particle in a controlled electromagnetic field. We assumed that the electric field is manipulable, and we set $\mathbf{E} = \mathbf{v}(k- \mid \mathbf{v} \mid)$. Obviously, and according to the Lyapunov theory, other velocity surfaces are possible by changing the term $\mid \mathbf{v} \mid$, for instance, to an ellipsoid, if $\mid \mathbf{v} \mid$ is replaced by $\sqrt{\frac{v_x^2}{a^2} + \frac{v_y^2}{b^2} + \frac{v_z^2}{c^2}}$, and so far.

Acknowledgments

This work is partially supported by the Spanish Ministry of Economy and Competitiveness through grant DPI2015-64170-R.

References

[1] S. Siboni (2008). Am J. Phys. 76, 649

[2] Y. Orlov, L. Aguilar, L. Acho (2008). Int. J. of Cntrol 81(2), 227

[3] R. L. Borelli, C. S. Coleman Differential equations: a modelling perspective, (John Wiley and Sons, Inc., New York, 1998) Chap. 9

Neutrino Oscillation and Neutrino Mass

S. Sahoo[1] and M. Goswami[2]

[1]Department of Physics, National Institute of Technology,
Durgapur – 713209, West Bengal, India. E-mail: sukadevsahoo@yahoo.com
[2]Department of Physics, Regional Institute of Education (NCERT),
Bhubaneswar - 751022, Orissa, India.

(*Submitted* 28-10-2015)

Abstract

In the standard model (SM) of particle physics neutrinos are massless and chargeless spin ½ particles. But the discovery of neutrino oscillations has shown that neutrinos have mass. From neutrino oscillations, we know only differences of mass-squared but not the absolute masses of individual neutrinos. In this article, we discuss neutrino oscillation and neutrino mass briefly.

1. Introduction

According to the standard model (SM) of particle physics fundamental constituents of matter are of two types: quarks and leptons [1]. This model assumes three generations (or families) of quarks and three generations of leptons (Table 1). Quarks are called (up, down), (charm, strange), and (top, bottom). The leptons consist of three flavours of charged leptons, the electron e^-, muon μ^- and tau τ^-, together with three flavours of neutrinos – the electron neutrino ν_e, muon neutrino ν_μ and tau neutrino ν_τ. All neutrinos are assumed to be massless and neutral.

Neutrinos are the second most abundant particles in the universe (photons are first). Neutrinos are very elusive and hardly interact with matter [2–7]. They do not enjoy electromagnetic and strong interactions but take part only in the weak interactions. They only interact with charged fermions and massive gauge bosons through the weak interaction. Neutrinos are copiously produced in the sun, in cosmic rays and even in laboratories. They are produced via the following processes:

(a) $(\nu_e, \bar{\nu}_e)$: Beta decay $(\bar{\nu}_e)$, Fission $(\bar{\nu}_e)$ and Fusion (ν_e) reactions.

(b) $(\nu_\mu, \bar{\nu}_\mu)$: Pion decay $(\pi^+ \to \mu^+ + \nu_\mu$ or the charge conjugate process).

(c) $(\nu_\mu, \bar{\nu}_e, \nu_e, \bar{\nu}_\mu)$: Muon decay $(\mu^- \to \nu_\mu + e^- + \bar{\nu}_e$ or the charge conjugate process).

Solar neutrinos are produced through process (a) while atmospheric (i.e. cosmic ray) neutrinos come from (b) and (c). Accelerator neutrinos rely on (b); reactor antineutrinos result from fission reactions (a). There are other neutrino sources e.g. supernovae etc. Physicists detected the first neutrinos from a supernova in 1987 when a star collapsed some 150,000 light-years away in the Large Magellanic Cloud, the galaxy nearest to the Milky Way.

Table 1: Three generations of Leptons and Quarks

First Generation	Second Generation	Third Generation
ν_e e^-	ν_μ μ^-	ν_τ τ^-
u d	c s	t b

2. Neutrino Puzzles and Neutrino Oscillation

In the 1960s, John Bahcall was trying to calculate what types of nuclear processes are occurring in solar fusion [2]. He predicted that the reaction $H^+ + H^+ \rightarrow He^{2+} + \nu + (\text{other})$ generates around 7×10^{10} neutrinos/(cm^2.s) on earth. Around 100 billion solar neutrinos are passing through our body every second. But they interact so weakly with other matter that remarkably little is known about them. The fusion reactions that take place in the sun only produce electron neutrinos. In order to detect these neutrinos he teamed up with experimentalist Ray Davis. Ray Davis and his team built a tank to hold 380,000 litres of perchloroethylene in the Homestake Gold Mine in South Dakota. They detected the solar electron neutrino flux at earth which was about 1/3 of the theoretical value. This was known as "*solar neutrino puzzle*". A similar discrepancy was also seen in atmospheric neutrinos. Atmospheric neutrinos are created as a consequence of cosmic ray protons from space hitting earth's atmosphere (which contains protons and neutrons). High energy proton/proton or proton/neutron collisions produce charged pions. These charged pions decay into muons and muon neutrinos. Then muons decay into an electron, an electron neutrino and a muon neutrino. Thus atmospheric neutrinos predict that for every electron neutrino there should be two muon neutrinos. But from IMB and Kamiokande experiments it was observed a ratio of one to one. This was "*atmospheric neutrino puzzle*".

In 1996, the SuperKamiokande detector was built in a zinc mine under 1,000 meters of solid rock in Japan. It was filled with 50,000 tons of ultra-pure water (not heavy water) and was designed to detect atmospheric neutrinos [7]. These neutrinos interact with atomic nuclei in the water to produce electrons, muons or tau leptons. Atmospheric neutrinos are mostly muon neutrinos. In 1998 [8], SuperKamiokande collaboration discovered that muon neutrinos converted or oscillated to tau neutrinos as they passed through the earth. The SuperKamiokande collaboration announced the first evidence for neutrino mass. Neutrinos oscillate in flavour because they have mass [9,10]. The SuperKamiokande was also used to study solar neutrinos. The fusion reactions that take place in the sun only produce electron neutrinos. But these neutrinos can subsequently oscillate into both muon neutrinos and tau neutrinos. Though the experiment was able to detect the solar neutrinos, it was unable to distinguish different neutrino types. Meanwhile, the Sudbury Neutrino Observatory (SNO) was constructed in a nickel mine under more than 2,000 meters of rock in Canada. Its tank was filled with 1,000 tons of heavy water. It was designed to study solar neutrinos. The SNO [11,12] could identify the electron neutrinos because it was filled with 'heavy water', which contains hydrogen nuclei with an extra neutron. The combined data from SuperKamiokande and SNO determined how many muon neutrinos or tau neutrinos were incident at the detector. The SNO results also provided further evidence for neutrino mass and confirmed that the total number of neutrinos from the sun agreed with theoretical calculations.

Takaaki Kajita was the team leader of the SuperKamiokande collaboration and Arthur B. McDonald directed the Sudbury Neutrino Observatory. On 6th October 2015, the Royal Swedish Academy of Sciences has announced to award the Nobel Prize in physics for 2015 jointly to *Takaaki Kajita*, University of Tokyo, Kashiwa, Japan and *Arthur B. McDonald*, Queen's University, Kingston, Canada "*for the discovery of neutrino oscillations, which shows that neutrinos have mass*". Their work has been published in an international reputed journal – Physical Review Letters [8,11,12].

3. Quantum Mechanics of Neutrino Oscillation

B. Pontecorvo [13] in 1958 and Z. Maki, M. Nakagawa and S. Sakata [14] in 1962 proposed that neutrino oscillation is a quantum mechanical effect. Neutrinos must have some mass for oscillations to occur. If neutrinos have finite masses, each flavour eigenstate (ν_e, ν_μ and ν_τ) can be expressed by a combination of mass eigenstates ν_1, ν_2 and ν_3 with mass m_1, m_2 and m_3 [15–17]. For simplicity let us discuss two flavour neutrino oscillation for example, between ν_e and ν_μ. Let $\theta_{e\mu}$ be the $\nu_e - \nu_\mu$ mixing angle. If a neutrino were produced as an ν_e at the source and travelled a distance L, the probability that it oscillated into a ν_μ is given by:

$$P(\nu_e \to \nu_\mu) = \sin^2 2\theta_{e\mu} . \sin^2\left(\frac{1.27\,\Delta m^2 . L}{E_\nu}\right) = A . \sin^2\left(\frac{\pi L}{\lambda}\right), \qquad (1)$$

where $\lambda = \dfrac{\pi E_\nu}{1.27\Delta m^2}$ acts as an oscillation length and $A = \sin^2 2\theta_{e\mu}$ as the amplitude of oscillation; $\Delta m^2 = \left| m_2^2 - m_1^2 \right|$ in eV^2 measures the difference of mass squares between the neutrinos, E_ν is the neutrino energy in GeV, and L in km is the distance of the detector from the neutrino source. Basically there are two variables: Δm^2 and θ. From equation (1) it is clear that

(a) The ideal distance of the detector from the source for observing the oscillations is $L = \lambda/2$, so that $\sin^2\left(\dfrac{\pi L}{\lambda}\right) = 1$.

(b) Δm^2 is dependent on (E/L); for small values of Δm^2 one needs small values for (E/L) to see the oscillations.

Here, we discuss oscillation between only two neutrino flavours ν_e and ν_μ. These flavour states can be expressed as the superposition of the mass eigenstates ν_1 and ν_2:

$$\begin{aligned} |\nu_e\rangle &= |\nu_1\rangle c + |\nu_2\rangle s \quad ; \\ |\nu_\mu\rangle &= -|\nu_1\rangle s + |\nu_2\rangle c \end{aligned}, \qquad (2)$$

where $c = \cos\theta$ and $s = \sin\theta$. For two flavours a single angle, θ, suffices to completely specify one basis in terms of the other. Consider now the state vector of an ν_e produced at t = 0. Thus, initially $|\psi(0)\rangle = |\psi_{\nu_e}\rangle = c|\psi_1\rangle + s|\psi_2\rangle$. (3)

If the stationary states $|\psi_1\rangle$ and $|\psi_2\rangle$ correspond to energies E_1 and E_2 respectively, then at a later time the state vector will be:

$$|\psi(t)\rangle = c|\psi_1\rangle \exp(-iE_1t) + s|\psi_2\rangle \exp(-iE_2t) \quad . \qquad (4)$$

The probability, $P(\nu_e, 0; \nu_\mu, t)$ of the state $|\psi(t)\rangle$ (originating as a ν_e at t = 0) appearing as a ν_μ is $\left|\langle\psi_{\nu_\mu}|\psi(t)\rangle\right|^2$ and is seen to be:

$$P(\nu_e, 0; \nu_\mu, t) = c^2 s^2 \left|-\exp(-iE_1t) + \exp(-iE_2t)\right|^2 \quad . \qquad (5)$$

The neutrinos are expected to have small masses, m_i, and are in the ultrarelativistic regime $\left(E_i \cong p + \frac{m_i^2}{2p}\right)$ where p (>> m_i) is the magnitude of the neutrino momentum. In this situation, we can rewrite the probability given in eq. (1):

$$P(\nu_e, 0; \nu_\mu, t) = \sin^2 2\theta_{e\mu} \sin^2\left(\frac{\pi L}{\lambda}\right), \qquad (6)$$

In the right hand side, the first factor is a consequence of the "mixing" while the second factor leads to the "oscillatory" behaviour. For vacuum oscillations, the former, dependent on the mixing angle $\theta_{e\mu}$, is a constant but in the Mikheyev-Smirnov-Wolfenstein (MSW) [18] matter effect, it changes with the matter density. From eq. (6)

$$P(\nu_e, 0; \nu_e, t) = 1 - P(\nu_e, 0; \nu_\mu, t)$$

$$= 1 - \sin^2 2\theta_{e\mu} \sin^2\left(\frac{\pi L}{\lambda}\right). \qquad (7)$$

From eq. (7) it is seen that $P(\nu_e, 0; \nu_e, t)$ is less than unity. The essential ingredients for this are twofold:

(i) The neutrinos must be massive and non-degenerate ($\Delta m^2 \neq 0 \Rightarrow \lambda$ is finite).

(ii) The mass eigenstates of the neutrinos ν_1, ν_2 must be different from the flavour eigenstates ν_e, ν_μ ($\sin 2\theta_{e\mu} \neq 0$).

An important theme of neutrino flavour change is the MSW effect, which is a matter-enhanced neutrino oscillation; in this case, the conversion $\nu_e \to \nu_x$ results from interaction between ν_e and solar electrons as the neutrinos travel from the centre of the sun. This effect originates from the additional interactions of a neutrino in a medium. It is well-known that interactions increase the inertia and mass is a measure of inertia. Thus, interactions in a medium result in a varying neutrino mass. While the solar neutrinos produced in the interior are coming out they pass through dense regions of the sun and experience the MSW effect. They interact with the solar electrons $\left(\nu + e^- \to \nu + e^-\right)$ giving to a contribution to the effective mass. From eq. (1) it is clear that the oscillation in flavour depends not on the mass of any particular neutrino type, but rather on the mass-squared difference between the flavours.

4. Neutrino Mass

The origin of neutrino masses is one of the biggest puzzles in particle physics today [6,7,19,20]. Although there are strong evidences for neutrino masses, till date we do not know the mechanism responsible for the generation of neutrino masses. Experiments have determined that neutrinos ν_1 and ν_2 have similar mass with ν_1 being lighter than ν_2. But till today we do not know whether ν_3 is much higher in mass ("normal hierarchy") or much lower in mass ("inverted hierarchy"). The absolute masses of neutrinos are not known, but a wide variety of experiments and theoretical models are setting their limits. A few of them are discussed below.

(i) Beta decay:

A nucleus with an overabundance of neutrons can transform to a more stable nucleus by emitting an electron and an antineutrino. This kind of process is known as beta (β) decay. The mass of the neutrino can be determined from the endpoint of the β-spectrum. The upper limit on the absolute scale of the electron neutrino mass is obtained from the tritium beta decay [21] as $m(\nu_e) < 2$ eV.

(ii) Neutrinoless double beta decay:

Double beta decay $(\beta\beta)$ is a nuclear transition (Z, A) $\rightarrow$ (Z+2, A) in which two neutrons bound in a nucleus are simultaneously transformed into two protons plus two electrons (there may be some light particles also). (i) In the two-neutrino double beta decay mode $(2\nu\beta\beta)$, there are $2\bar{\nu}_e$ emitted together with $2e^-$. The lepton number is conserved for this mode and this mode of decay is allowed in the standard model of electroweak interaction. (ii) In the neutrinoless double beta decay mode $(0\nu\beta\beta)$, only the $2e^-$ are emitted and nothing else. This neutrinoless double beta decay occurs when the two antineutrinos, instead of manifesting themselves as real states, "annihilate". This can only occur if neutrinos are their own antiparticles. This mode violates the law of lepton number conservation and is forbidden in the standard model. Hence its observation may lead to a signal of "new physics". The lepton number violation can generate a lepton asymmetry in the early universe, which will be able to explain the present baryon asymmetry of the universe.

Neutrinoless double beta decay is the only experiment that can probe the Majorana nature of the neutrino (i.e. the neutrino and antineutrino are identical) [22–25]. The values of the neutrino mass-squared differences are known, but the absolute values of neutrino masses are elusive. The observation of neutrinoless double beta decay would not only reveal the neutrinos are Majorana fermions, but would also provide information regarding the absolute values of the neutrino masses. The two-neutrino double beta decay has already been experimentally observed. There is possible evidence of neutrinoless double beta decay in the Heidelberg-Moscow experiment [22] but so far, neutrinoless double beta decay has not yet been observed conclusively [25].

Assuming Majorana nature of neutrino, a strong limit on the mass eigenstate of ν_e is obtained as $m_{\nu 1} \leq 0.4 - 0.5$ eV from neutrinoless double beta decay experiments with Germanium [26,27] and Tellurium [22,28]. Furthermore, the search for the $0\nu\beta\beta$ decay is the only way to probe the Majorana nature of neutrinos and one of the most promising ways to search for lepton number violation.

(iii) Neutrino oscillations: In this method, neutrino mass squared differences $\Delta m_{ij}^2 = m_i^2 - m_j^2$ are determined. The two different Δm^2 values are $\left|\Delta m_{atm}^2\right| = (1.9 - 3.0) \times 10^{-3}\,\text{eV}^2$ and $\Delta m_{sol}^2 = 8.0^{+0.4}_{-0.3} \times 10^{-5}\,\text{eV}^2$. This range and indicated error bars show the present sensitivity. This mass determination is independent of the charge conjugation properties of neutrinos.

(iv) Cosmological observations: From cosmic microwave background and large scale structure data, the size of fluctuations is observed at different scales. Since the light neutrinos would have smeared out fluctuations at small scales, the power spectrum at small scales is sensitive to the neutrino mass. Although the absolute mass of the neutrinos have not yet been determined, there is an upper bound on the sum over all neutrino masses from cosmological observations [29]: $\sum_{i=e,\mu,\tau} m_{\nu_i} \leq 0.61$ eV, which are to some extent model- and analysis dependent [30]. This mass determination is independent of the Majorana or Dirac nature of neutrinos.

5. Conclusion

The electron-neutrinos produced from the sun were measured to be less than what was predicted by the standard solar model and experiments with the atmospheric neutrinos demonstrate that there was a depletion of atmospheric muon-neutrinos while there was no depletion of electron-neutrinos. One possible explanation for the observed solar neutrino deficit is that the ν_e produced in the centre of the sun could convert itself to another type, i.e., $\nu_e \rightarrow \nu_x$ with $x = \mu$ or τ, during its passage to the earth via a process called *neutrino oscillation*. Similarly, the atmospheric muon-neutrino deficit could be due to the conversion of ν_μ to ν_τ. In this way, "atmospheric neutrino puzzle" and "solar neutrino puzzle" were resolved by neutrino oscillations in 1998 and 2001(2) respectively. Δm^2_{23}, $\sin^2 2\theta_{23}$, Δm^2_{12} and $\sin^2 2\theta_{12}$ have been measured accurately by the present generation experiments assuming 2-flavour neutrino oscillations [15]. It is also possible that CP invariance can be violated in the lepton sector. Neutrino oscillation is a quantum mechanical effect. Neutrinos must have some mass for oscillations to occur. The Nobel Prize in physics for 2015 has been awarded jointly to *Takaaki Kajita*, University of Tokyo, Kashiwa, Japan and *Arthur B. McDonald*, Queen's University, Kingston, Canada "*for the discovery of neutrino oscillations, which shows that neutrinos have mass*". From neutrino oscillations, we know only differences of mass-squared but not their individual masses. Although the absolute masses of neutrinos are not known, a wide variety of experiments and theoretical models are setting their limits as discussed in section 4. Recently, Robertson [31] has discussed neutrino mass. According to him, neutrino oscillations set a lower limit of 0.02 eV and upper limit from measurements is 2.0 eV. Recently, Fritzsch [32] has calculated the masses of three neutrinos: $m_1 \approx 0.003$ eV, $m_2 \approx 0.012$ eV, and $m_1 \approx 0.048$ eV.

Neutrinos show up in precision cosmological observations: since they have a small mass they should cluster on sufficiently large scales. Neutrinos may also be messengers of dark matter annihilation in our galactic halo or in the core of the sun. Neutrinos are important for the study of the sun, stars, core-collapse supernovae, the origins of the cosmic rays, the large-scale structure of the universe, and big bang nucleosynthesis. These tiny neutrino masses are of great interest because they might arise from some fundamentally different mechanism to the way the masses of other particles are generated i.e. the Higgs mechanism. Although the SM is very successful to explain many low as well as high energy phenomena in particle physics but within the framework of this model it is not possible to realize the massive neutrinos. The existence of neutrino mass is one of the signatures of new physics beyond the SM [33–35].

There are still many things about neutrinos that we need to know. A few of them are: (i) absolute mass of neutrinos (ii) whether neutrinos are Majorana particles ($\bar{\nu}_i = \nu_i$) or Dirac particles ($\bar{\nu}_i \neq \nu_i$), (iii) What is the pattern of neutrino masses (normal mass hierarchy or inverted mass hierarchy)? (iv) Why neutrino masses are so small or why there is such a large gap between the neutrino and the charged fermion masses? (v) CP violation in neutrino (lepton) sector, etc. We hope the results from further experiments will provide us the answer to these problems in near future.

The Karlsruhe Tritium Neutrino Experiment (KATRIN) in Germany is taking data to make a very precise measurement of the electron energy spectrum from beta decay. We hope we can know the mass of neutrinos from KATRIN and astronomical surveys very soon. The current generation of oscillation experiments including Double Chooz, RENO, Daya Bay, T2K and NOvA, will try to resolve the neutrino mass hierarchy. From the ongoing and future neutrino experiments, we expect more surprises. Neutrinos have and will continue to provide

important information on structure formation in the early universe, earth, solar and supernova physics, nuclear properties, and rare decays of charged leptons and hadrons [36]. The study of neutrino physics and the implications of the results connect many disciplines together, from particle physics to nuclear physics to astrophysics to cosmology. Thus, neutrino physics continues to be a very exciting field and may also bring us new surprises in this 21st century.

Acknowledgment

We thank the referee for suggesting valuable improvements of the manuscript.

References

1. S. Adams, *Frontiers – Twentieth Century Physics*, Replika Press Pvt. Ltd., Delhi, pp. 168 – 189 (2000).
2. D. Lincoln and T. Miceli, *The Physics Teacher*, **53**, 331-338 (2015).
3. P. Langacker, J. Erler and E. Peinado, *J. Phys. Conf. Ser*. **18,** 154 – 187 (2005) [arXiv: hep-ph/0506257].
4. G. Altarelli, arXiv: hep-ph/ 0611117 (2006).
5. D. P. Roy, *Physics News–Bulletin of Indian Physics Association*, **39**(3), 51 (2009) [arXiv: 0809.1767 [hep-ph]].
6. S. F. King, *Contemp. Phys*. **48**(4), 195 – 211 (2007).
7. H. Murayama, *Phys. World*, **15**(5), 35–39 (2002).
8. Y. Fukuda, T. Kajita et al. [Super-Kamiokande Collaboration], *Phys. Rev. Lett.* **81**(8), 1562 (1998).
9. E. Hecht, *The Physics Teacher*, **41**, 164 (2003).
10. M. Goswami and S. Sahoo**,** *Science Reporter*, **45**(4), pp. 33 – 34 (2008).
11. Q. R. Ahmad, A. B. McDonald et al. [SNO Collaboration], *Phys. Rev. Lett.* **87**(7), 071301 (2001).
12. Q. R. Ahmad, A. B. McDonald et al. [SNO Collaboration], *Phys. Rev. Lett.* **89**(1), 011301 (2002).
13. B. Pontecorvo, *Sov. Phys. JETP* **6**, 429 (1958); **7**, 172 (1958).
14. Z. Maki, M. Nakazawa, S. Sakata, *Prog. Theor. Phys*. **28**, 870 (1962).
15. T. Kajita, *Int. J. Mod. Phys. A*, **24**, 3437–3446 (2009).
16. R. N. Mohapatra and A. Y. Smirnov, *Ann. Rev. Nucl. Part. Sci.* **56,** 569 (2006).
17. R. N. Mohapatra et al., *Rept. Prog. Phys.* **70,** 1757-1867 (2007)[arXiv:hep-ph/ 0510213].
18. M. Mikheyev, A. Smirnov, *Sov. J. Nucl. Phys*. **42**, 913 (1986); L. Wolfenstein, *Phys. Rev*. **D17**, 2369 (1978), ibid **D20**, 2634 (1979).
19. S. Sahoo, *Indian J. Pure & Appl. Phys.* **48**(10), 691–696 (2010).
20. S. Sahoo, A. Pathak and A. Mondal, *ARPN Journal of Science and Technology*, **4**(6), 345 – 348 (2014).
21. E. W. Otten and C. Weinheimer, *Rept. Prog. Phys*. **71**, 086201 (2008). [arXiv:0909.2104 [hep-ex]].
22. H. V. Klapdor-Kleingrothaus, A. Dietz, H. L. Harney and I. V. Krivosheina, *Mod. Phys. Lett. A* **16**, 2409 (2001) [hep-ph/0201231].
23. D. R. Artusa et al. (CUORE Collaboration), *Eur. Phys. J. C* **74**(10), 3096 (2014) [arXiv: 1404.4469 [nucl-ex]].
24. S. Sahoo and B. Sahoo, *IAPT Bulletin*, **2**(2), 37–41 (2010).
25. J. J. Gómez-Cadenas & Justo Martín-Albo, arXiv:1502.00581 [hep-ex] (2015).
26. C. Weinheimer et al., *Phys. Lett. B*, **460,** 219 (1999).
27. V. M. Lobashev et al., *Phys. Lett. B*, **460,** 227 (1999).
28. C. E. Aalseth et al. [IGEX Collaboration], *Phys. Rev. D***, 65,** 092007 (2002).
29. E. Komatsu et al. [WMAP Collaboration], *Astrophys. J. Suppl*. **180,** 330 (2009) [arXiv: 0803.0547].
30. S. Hannestad et al., *J. Cosmol. Astropart. Phys*. **08,** 015 (2007).

31. R. G. H. Robertson, *Nuclear and Particle Physics Proceedings*, **265–266**, 7–12 (2015) [arXiv: 1502.00144 [nucl-ex]].
32. H. Fritzsch, *Mod. Phys. Lett. A* **30**(16), 1530012 (2015) [arXiv: 1503.01857 [hep-ph]].
33. G. Senjanovic, arXiv: 0911.0029 [hep-ph] (2009).
34. I. Bhattacharya, *Fundamental Journal of Mathematical Physics,* **3,** 23–31 (2013) [arXiv: 0910.5801 [hep-ph]].
35. A. Y. Smirnov, arXiv: 0910.1778 [hep-ph] (2009).
36. A. de Gouvea et al., arXiv: 1310.4340 [hep-ex] (2013).

Physics Through Problem Solving XXIX

Problems in mechanics

S.N. Maitra

Professor (retd.) and Former Head of Maths Department
National Defence Academy
Pune 411 021.

E-mail: soumen_maitra@yahoo.co.in

Problem no 1:

A heavy uniform rectangular lamina is submerged vertically in still water. One corner of the lamina is pulled above water by a vertically suspended string so that its diagonal lies on the water surface as shown in figure 1.Find the specific gravity of the lamina

Solution to problem no 1:

Let 2a and 2b be the breadth and length of the lamina ABCD of weight W submerged in water, T be tension of the vertical taut string fastened at its corner A; BD is the diagonal lying on the water surface above which is the portion BAD and below which is the portion BCD of the lamina; G is the Center (middle point) of the lamina such that BG = GD.

Let O be the Center of pressure at which the force W1 of buoyancy acts vertically upwards. This is illustrated in figure 1. Obviously the lamina floats in still water under the action of forces:

1. Tension T of the string acting vertically upwards at A and passing through E on the horizontal line BD and through I on side BC

2. Weight W of the lamina acting vertically downwards at the mid point G of diagonal BD .

3. Force W1 of buoyancy as depicted above , pass through F on the diagonal BD and through the point O which is also the centroid of the portion BCD immersed in water. This is depicted in figure 1.

If ρ and ρ_0 be the densities of the lamina and water respectively and t the thickness , then we have

$$W = 2a.2bt\rho = 4\,abt\rho$$

W_1 = also weight of water of volume equal to that of the immersed portion BCD i.e. half the volume of the lamina = $2abt\rho_0$ (1)

For flotation i.e. equilibrium of the lamina, resolving the forces horizontally and vertically and taking moments of the forces about G_1, we get because of (1)

$$T = W - W_1 = 2abt\,(2\rho - \rho_0) \quad (2)$$

$$T.EG = W_1 . GF \quad (3)$$

Which in consequence of (1) and (2) becomes $2abt(2\rho - \rho_0)EG = 2abt\rho_0 .GF$

and hence the specific gravity of the lamina becomes $S = \rho/\rho_0 = \frac{1}{2}(GF/EG + 1)$ (4)

Now we need to find GF and EG by geometry.

In the rectangular lamina, AB = CD = 2a , BC = AD =2b so that diagonal

$BD = \sqrt{a^2 + b^2}$; $\angle A = \angle B = \angle C = \angle D = 90^0$

In the right angled triangle BDC , H being the midpoint of side CD ,

BH = $\sqrt{a^2 + 4b^2}$ and O being the centroid i.e. the Centre of pressure,

OH = BH/3 = $\sqrt{a^2+4b^2}/3$

OB = 2BH/3 = 2/3 .

$\sqrt{a^2 + 4b^2}$ (5)

BG = GD = (1/2) BD = $\sqrt{a^2+b^2}$

BE = ABcos∠ABD = 2acos∠CDB = 2acos∠CDB = $(2a^2)/ (\sqrt{a^2 + b^2})$ ($\angle B = \angle E = 90^0$)

EG = BG – BE = $\sqrt{a^2+ b^2} - 2a^2/(\sqrt{a^2+ b^2}) = (b^2-a^2)/(\sqrt{a^2 + b^2})$ (6)

Let ∠DBH = θ, ∠DBC = θ_1 , ∠CDH = θ_2, so that

$Cos(\theta_1 - \theta_2) = cos\theta_1 cos\theta_2 + sin\theta_1 sin\theta_2$

$= (b/(\sqrt{a^2 + b^2})).2b/(\sqrt{a^2 + b^2}) + a/(\sqrt{a^2 + b^2}).a/(\sqrt{a^2 + 4b^2})$

Or,

$Cos\theta = cos(\theta_1 - \theta_2) = a^2 +2b^2/(\sqrt{a^2 + b^2})(\sqrt{a^2+4b^2})$ (7)

By (5), BF = OB cosθ = (⅔) $\sqrt{a^2 + 4b^2}(a^2 + 2b^2)/((\sqrt{a^2 + b^2})(\sqrt{a^2 +4b^2})) = ⅔(a^2 +2 b^2)/\sqrt{a^2 + b^2}$

$\angle F = 90^0$

GF = BF – BG = (⅔)$(a^2 + 2b^2)/(\sqrt{a^2 + b^2}) - (\sqrt{a^2 + b^2})$ $= (b^2 - a^2)/3(\sqrt{a^2 - b^2})$ (8)

Using (6) and (8) in (4) we get

$S = ½((b^2 - a^2)/3(b^2 - a^2) + 1) = ⅔$ $(b \neq a)$ (9)

which is independent of the dimensions of rectangular lamina. This shows that if any rectangular lamina is submerged in a liquid in this manner, its density will be two third the density of the liquid.

DISCUSSION : If the density of water $\rho_0 = 1$, from (9) the density of the lamina ⅔ gm/c.c. The foregoing finding reveals that without the string the lamina can float with its one third portion above the water. From (1) and (2) one can get the weight of the lamina and tension of the string as

W = 8abt / 3 and

T = 2abt/3

Which shows that the tension is ¼ th times the weight of the lamina.

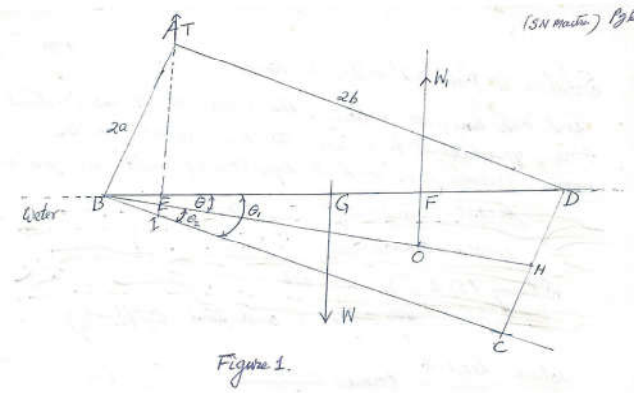
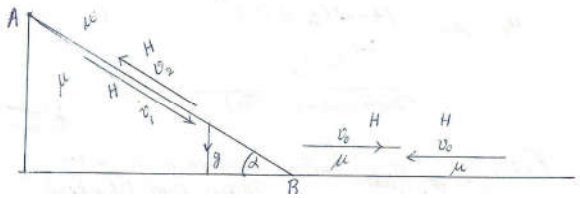


Figure 1.



Problem No. 2:

A motor car undertakes a trip downhill, on level and then returns to same spot along the same route. It travels downhill, on level and uphill with speeds v_1, v_0 and v_2 respectively with the same constant horsepower. Consider the hill as inclined plane and the road resistance with the same co-efficient of friction. Find the inclination of the hill and the coefficient of friction.

Solution to problem no 2:

Let μ and α be the coefficient of friction of the road along the hill and on the level, and the inclination of the hill to the horizontal , g the acceleration due to gravity and m the mass of the car. By the definition of Horsepower H, (vide figure 2)

H = applied force x velocity = (mgμ cosα - mgsinα) x v_1 = mgμv_0 (1)

= (mgμcosα + mgsinα) x v_2 = H.P

While traveling on the hill of inclination α to the horizontal, the normal reaction on the car is

R = mgcosα

and the frictional force on it is

F = μR = mgμcosα

Where as the gravitational force on it is

G = mgsinα

When it moves on level ground, normal reaction = mg and the frictional force = μmg When the car travels with constant horsepower and with uniform velocity, the resultant of the frictional force, gravitational force and the force applied by the engine becomes zero and its equations of motion are given (1)

μcosα – sinα = μv_0/v_1 (2)

μcosα + sinα = μv_0/v_2 (3)

Solving (2) and (3) one gets

μcosα =(μv_0/2) (1/v_1 + 1/v_2) and sinα = (μv_0/2)(1/v_2 - 1/v_1)

Which lead to tanα = μ(v_1 - v_2)/(v_1 + v_2) (4)

Or , √ (4 - v_0^2(1/v_1 + 1/v_2)^2) / v_0(1/v_1 - 1 /v_2)