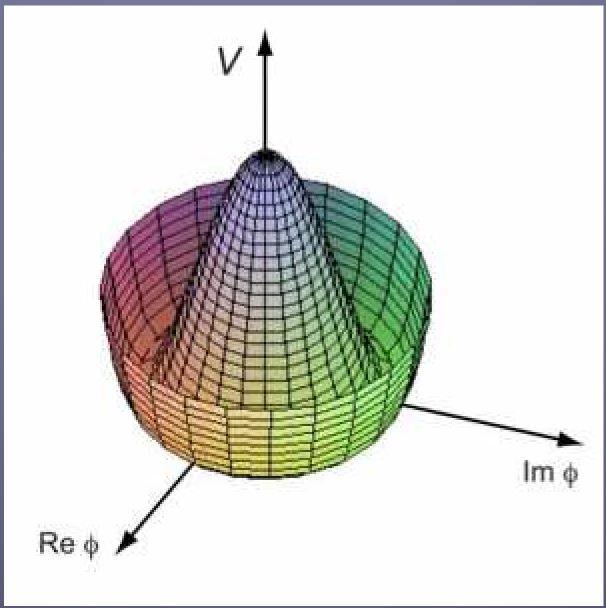
Vol 28 No. 3 Jul - Sep 2012 ISSN 0970-5953

INDIAN ASSOCIATION OF PHYSICS TEACHERS

IAPT

PHYSICS EDUCATION



V (Φ) vs Φ illustrating degerate vacuum states

www.physedu.in

Volume 28, Number 3

In this Issue

EDITORIAL

(*Submitted* 25 - 09 - 2012)

To begin with, I throw up a question to ponder upon, which concerns all of us. I do not have the exact demographic data, but the number of Indians with age between 5 years and 25 years is easily estimated to be about forty crores. How do we get these young people to be educated acquiring variety of skills and knowledge required to be useful to the society? To escalate the challenge, this has to be achieved within a span of five to ten years, otherwise we lose a generation. Whatever we do, this goal simply cannot be reached by recruiting more and more teachers to teach in classrooms. (Where are we going to get the teachers from?) A large part of the solution lies in building up the facility and culture of distance education. By employing appropriate technologies, it is possible to reach nook and corner of our country and the domain expertise of the available teachers can be used to make an impact on enormously large number of students as compared to class room teaching. The distance education may not completely make up for every educational requirement like laboratory training in physics and engineering or treating actual patients in medicine and surgery. These requirements need to be dealt with separately. In a nutshell a systematic planning and development of the field of distance education is inevitable in a country like India, where young people are looked upon as the major resource to its prosperity. An effective development and implementation of the infrastructure and the eduware for this program is a challenge that we have to take and succeed in it. In view of this, Physics Education is now an on-line journal, can be accessed from anywhere and will soon support multi-media inputs, so as to become an effective platform for distance education. We certainly welcome contributions involving distance education techniques like simulations of processes and from people who would like to describe their experiences in developing and using courses for distance education.

Coming to this particular issue, I find it pretty interesting. I must mention the article by Archimann Raju, a junior college student, on the brachistochrone problem with resistance. The algorithm presented by him is of general applicability and can be actually used in practice. A similar kind of work dealing with Snell-Descartes law for massive particles is presented by D.N.Basu. S.Shivakumar's article answers a long standing question in students' mind: why are complex numbers needed in quantum mechanics? Although this question is succinctly dealt with in quantum mechanics text books like Modern Quantum Mechanics by Sakurai, we feel that this article will make the salient features clear to a large class of students, especially those who do not have an easy access to books like Sakurai. The regular features like 'Physics Through Problem Solving' by Ahamad Sayeed and 'Physics through Laboratory' (the article on compound pendulum by Pathare et al.) are educative as always. The paper by Sanjay Harrison and Sindhu Vincent is an example of a quick estimate of the gas pressure in a balloon using very simple experiments. Finally, in his article on Higgs boson Prof. Ramachandran eliminates the myth that Higgs boson is responsible for creation of mass in all situations.

I wish you a happy reading.

Pramod S. Joag.
Chief-editor@physedu.in,
pramod@physics.unipune.ac.in
Chief Editor

A simple way to solve the brachistochrone problem with resistance

Archishman Raju[1]

[1]Department of Physics
St. Stephen's College
Delhi 110007, India.
archishmanraju@hotmail.com

(Submitted 04-07-2012)

Abstract

Variational problems are ubiquitous in physics. But an introductory course on the calculus of variations is typically restricted to solving a few standard problems like the classical brachistochrone. Several experiments have clearly shown this theory to be inadequate, because any actual physical situation involves resistance, but no attempt has been made so far to reconcile experiment with theory. Adding resistive forces to the problem makes analytical solutions intractable. We show how such hard variational problems can be easily solved using a simple numerical approach. This allows a large variety of variational problems to be solved at an introductory level and the solution checked against simple experiments. We illustrate this by solving the brachistochrone problem with Coulomb friction and fluid resistance. We outline an experiment which could be used to check the result.

1 Introduction

Variational principles are ubiquitous in physics. Yet an introductory course on the calculus of variations treats very few problems. There is also no simple numerical method (as with ordinary differential equations) through which a larger variety of problems can be examined. The classical brachistochrone problem is the standard problem solved in introductory textbooks [1]. However, the addition of any kind of resistance makes the problem much harder to solve. The addition of Coulomb friction was first examined in [2] and requires a constrained variational technique. An examination of [3] reveals that simple models of fluid resistance lead to very involved algebra.

There have been several experiments on the brachistochrone performed at the undergraduate/high school level and all of them have shown significant deviations from the expected result. Moreover, because the theory with resistance is too complicated, there has been no attempt to incorporate it. As an example, take the isochronous property of the cycloid which is also the brachistochrone without resistance. Introductory physics courses teach that the simple pendulum has an amplitude dependent time period which makes it unsuitable as a clock. The cycloidal pendulum is proposed as the solution to this problem on the belief that it is isochronous. In a recent experiment with high school students however [4], it was found that that a real cycloidal path is not truly isochronous and a definite amplitude-dependence was observed, as is to be expected (however, no theoretical examination was attempted). The obvious cause of the experimental deviations is that a real cycloidal pendulum (as opposed to an idealised one) involves resistance.

Similarly, an experiment with 'Hot-Wheels' cars found the cycloid to be the fastest path among those that were tried, but the authors did not examine the difference in the experimental and theoretical time [5]. Another experiment with undergraduates again found significant difference in the theoretical and experimental time values [6] but did not investigate the possibility that the cycloid is no longer the brachistochrone when friction is included.

We present a simple numerical method which can be used to solve any variation of the problem. In particular, it can be used to *quantitatively* examine how resistive forces affect the solution and hence obtain agreement between theory and experiment. Using this, one can even ask more complicated questions, like, what is the shortest path underwater? Is it still a cycloid? This, too, can be directly linked to a simple experiment (as we explain later).

The brachistochrone problem with friction has been considered by other authors [7, 8, 9, 10]. The numerical approach found in these references is mostly limited to obtaining numerical solutions to the Euler equation. Numerical solutions to partial differential equations are well known. The real difficulty is to first formulate these problems variationally. To get over this difficulty, we use the fact that these situations are simple from a Newtonian point of view. This makes it easy enough to be used in introductory courses.

In a different context, there is a numerical approach to variational problems in mechanics [11]. The algorithm used in [11] does not directly apply to our problem because we seek to minimize the time of descent in the presence of non-conservative forces. Further, our numerical algorithm is useful not only as an educational tool, but can also be applied to solve a wide variety of variational problems where the analytical solution is not feasible. We demonstrate this by solving the brachistochrone problem with fluid resistance and Coulomb friction.

2 Algorithm

The mathematical problem at hand becomes much easier if seen as a physical problem (like a bead sliding down a pipe filled with water). We begin by discretizing the x axis into some N points. Specifying the y values at those points completely defines the curve. We now need to minimize the time it takes for the bead to slide down. It is simple to formulate this from a Newtonian point of view. We need to minimize the time it takes for the bead to travel down a path. The time of travel is obtained by solving the equations of motion. While in most cases, an analytical solution will not exist, it is easy to solve the equations of motion numerically.

To calculate the time in this way we need, first, to construct a path, given the y coordinates at the N points. One could use a straight line between the points. However, this is not a very good choice for the present problem from a numerical and algorithmic point of view because the lines do not join smoothly, and differentiability fails at those points. A better choice is a smoothened polynomial. Let us for the moment say that the path is given by a function $y(x)$. The forces involved are the force of gravity, the buoyant force, Coulomb friction and fluid resistance. For the fluid resistance, we assume a resistance proportional to the square of the velocity. The coefficient will depend on the nature of the fluid and the object. Then the equations of motion for the system are

$$m\ddot{x} = mg_e \sin\theta \cos\theta - \mu m g_e \cos^2\theta - kv\dot{x}, \tag{1}$$

$$m\ddot{y} = -mg_e \sin^2\theta + \mu m g_e \cos\theta \sin\theta - kv\dot{y}, \tag{2}$$

$$\theta = -\tan^{-1} f'(x), \tag{3}$$

$$g_e = \frac{(m - \frac{4}{3}\pi\rho r^3)g}{m}. \tag{4}$$

Here m is the mass, ρ is the density of the fluid, μ is the coefficient of friction, k is the drag coefficient and $v = \sqrt{\dot{x}^2 + \dot{y}^2}$. The dot denotes derivative with respect to time whereas the prime denotes derivative with respect to x. g is the acceleration due to gravity whereas g_e is the effective acceleration due to gravity when the buoyant force is taken into account. The buoyant force has been calculated for a spherical object.

Let us say we want to find the brachistochrone between points $(0,\ 0)$ and (a, b). The algorithm starts by calculating the time it takes to cover this distance through some initial path (like a straight line). This path is defined by a set of N points (x_i, y_i) between $(0, 0)$ and (a, b), where the x_i points are taken to be fixed. Now, the algorithm proceeds by sequentially updating the y_i points by changing them by a specified small amount. So it starts by increasing (decreasing) y_1. This gives a new curve $y(x)$. The equations of motion are solved again to obtain a new time of descent. If this time of descent is smaller than the previous one, y_1 is increased (decreased) again. This is continued till a change in y_1 leads to an increase in the time of descent. Then the algorithm proceeds to y_2 and repeats the same process. After reaching y_N the algorithm comes back and updates y_1 again. It stops when no possible step leads to a decrease in time.

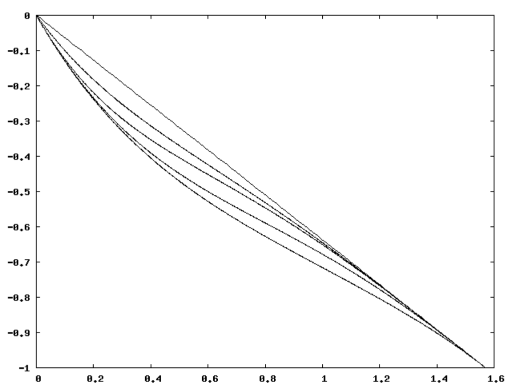


Figure 1: The plot shows typical curves the algorithm tries before reaching the solution.

3 Results

For the present problem, we use a Bezier curve to interpolate between the points. Even though we are solving 2 equations of motion, the curve is actually defined by 1 parameter, say t. We use a Runge-Kutta 4th order solver to solve the equations of motion. Though the equations of motion implicitly constrain the object to move along the curve, it is possible for numerical errors to develop. Hence, for each step the ODE solver takes, rather than computing the derivative of the curve (which is required in the equations of motion) from either the current x or y position of the object, the x, y position is first mapped to t using a simple linear search. The derivative of the curve is then calculated at point t. We show typical steps in the algorithm in Figure 1. If needed, a more sophisticated optimization algorithm can also easily be applied to the problem as formulated. As an example, the simulated annealing algorithm can be used since the problem has a cost function as well as a specified way to change its state. We show simulation results for different values of $\frac{k}{m}$ in Figure 2. The least time curves obtained are between the cycloid and the straight line. As the drag coefficient increases, the curves start resembling a straight line. For a large enough drag coefficient, the least time curve is the straight line.

To check this independently, it is possible to perform a simple experiment for a fluid (say, water) which we briefly describe. This experiment requires only a flexible pipe filled with water. By fastening the pipe at some appropriate points it can be made to resemble a smooth curve passing through those points. Even though the exact shape of the pipe might be difficult to ascertain, the shape of the pipe can be approximated by a smoothened polynomial through those points. The time a ball bearing takes to slide down the pipe can be measured and hence the time it takes to slide down different curves can be experimentally compared.

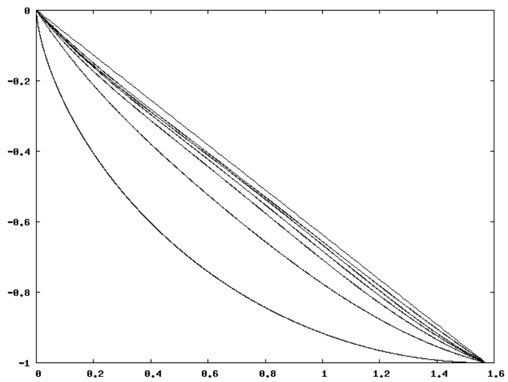


Figure 2: The plot shows the brachistochrone between points (0, 0) and ($\frac{\pi}{2}$, -1). The effective gravity because of the buoyant force is 8.72 m/s^2. We have $\mu = 0.1$. Starting from above the plots are that of a straight line, the curve obtained for $\frac{k}{m} = 11$, $\frac{k}{m} = 7$, $\frac{k}{m} = 5$, $\frac{k}{m} = 3$, $\frac{k}{m} = 1$ and the cycloid between the two points.

4 Concluding remarks

The brachistochrone problem has earlier been suggested to be the best introduction to variational calculus. On this note, an earlier project with undergraduates tried to analyze this problem in detail using both theory and experiment [6]. The difference in theory and experiment (due to resistance) could not be addressed since their numerical method was limited to evaluating the time integral of the classical brachistochrone problem for different curves. We have shown how this problem, which is hard even to formulate analytically, from a variational point of view, can be easily solved using a simple numerical scheme. Moreover, the solution can be checked with experiments easy enough to perform in the classroom.

References

[1] G F Simmons. *Differential Equations.* Tata McGraw-Hill, 1974.

[2] N Ashby etal. Brachistochrone with Coulomb friction. *American Journal of Physics*, 43(10):902–906, October 1975.

[3] B Vratanar and M Saje. On the analytical solution of the brachistochrone problem in a non-conservative field. *International Journal of Non Linear Mechanics*, 33(3):489–505, 1998.

[4] P Gluck. Motion on cycloid paths: a project. *Physics Education*, 45:270, 2010.

[5] F M Phelps III, F M Phelps IV, B Zorn, and J Gormley. An experimental study of the brachistochrone. *European Journal of Physics*, 3, 1982.

[6] M Desaix, D Anderson, and M Lisak. The brachistochrone problem-an introduction to variational calculus for undergraduate students. *European Journal of Physics*, 26:857, 2005.

[7] V Čović and M Vesković. Brachistochrone on a surface with Coulomb friction. *International Journal of Non-Linear Mechanics*, 43(5):437–450, 2008.

[8] J C Hayen. Brachistochrone with Coulomb Friction. *International Journal of Non-Linear Mechanics*, 40:1057–1075, 2005.

[9] A S Parnovsky. Some generalisations of brachistochrone problem. *Acta physica Polonica. A*, 93, 1998.

[10] O Jeremić, S Salinić, A Obradović, and Z Mitrović. On the brachistochrone of a variable mass particle in general force fields. *Mathematical and Computer Modelling*, 54:2900–2912, 2011.

[11] S Tuleja, E F Taylor, and J Hanc. Use of computer in advanced mechanics - principle of least action. *American Journal of Physics*, 71(4):386, 2003.

On the E-L equation and Snell's law for massive particles : a mathematical revisit

D. N. Basu

Variable Energy Cyclotron Centre
1/AF Bidhan Nagar
Kolkata 700 064, India.
dnb@vecc.gov.in

(Submitted 30-04-2012)

Abstract

This article describes: (i) the conditions to check whether Euler-Lagrange equation for extremisation provides minimum or maximum and (ii) the derivation of Snell-Descartes law for massive particles which is in contradiction to that for the light waves.

Keywords: Euler-Lagrange equation; Snell-Descartes law.

1 On the Euler-Lagrange equation

In calculus of variations [1], the Euler-Lagrange equation, Euler's equation or Lagrange's equation, is a differential equation whose solutions are the functions for which a given functional is stationary. The Euler-Lagrange equation is useful for solving optimization problems in which, given some functional, one seeks the function minimizing (or maximizing) it because a differentiable functional is stationary at its local maxima and minima. In Lagrangian mechanics, because of Hamilton's principle of stationary action, the evolution of a physical system is described by the solutions to the Euler-Lagrange equation for the action of the system. In classical mechanics, it is equivalent to Newton's laws of motion, but it has the advantage that it takes the same form in any system of generalized coordinates, and it is better suited to generalizations. However, in other practical problems one, very often, needs to know whether the solutions of Euler-Lagrange equation provides local maxima or minima. In this short note, the condition will be obtained that determines whether these solutions represent local maxima or minima.

Let us consider a fixed end point problem defined by the integral

$$I = \int_{x_1}^{x_2} f(y', y, x)dx \quad (1)$$

where x_1 and x_2 are the fixed end points and $f(y', y, x)$ is an explicit function of the $y' = \frac{dy}{dx}$, $y = y(x)$ and x. The problem is now of extremising the functional I. For this let us make variations as

$$y \to y + \alpha\eta, \quad y' \to y' + \alpha\eta' \quad (2)$$

where $\eta = \eta(x)$, $\eta' = \frac{d\eta}{dx}$ and α is independent of x and demand that the functional I is stationary under such variations brought about by the parameter α such that the functional I attains an extremum at $\alpha = 0$ that is $\frac{\partial I}{\partial \alpha}|_{\alpha=0}$. Taylor's expansion up to first order yields

$$f(y'+\alpha\eta', y+\alpha\eta, x) = f(y', y, x)+\alpha\eta f_y+\alpha\eta' f_{y'} \quad (3)$$

where $f_y = \frac{\partial f}{\partial y}$ and $f_{y'} = \frac{\partial f}{\partial y'}$. The condition for extremum can be obtained using the expansion of Eq.(3)in Eq.(1) to provide:

$$\frac{\partial I}{\partial \alpha}|_{\alpha=0} = 0 = \int_{x_1}^{x_2} [\eta f_y + \eta' f_{y'}]dx. \quad (4)$$

The second term on the right hand side of the above equation can be integrated by parts as

$$\int_{x_1}^{x_2} \eta' f_{y'} dx = [\eta f_{y'}]_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{df_{y'}}{dx}\eta dx. \quad (5)$$

Since it is a fixed end point problem, the values of $y(x)$ should not vary at the end points x_1 and x_2 implying $\eta(x_1) = 0 = \eta(x_2)$ which makes the first term on the right hand side of Eq.(5) to vanish and one obtains

$$\int_{x_1}^{x_2} [f_y - \frac{df_{y'}}{dx}]\eta dx = 0. \quad (6)$$

Since the function $\eta(x)$ is quite arbitrary, yields the Euler-Lagrange equation:

$$f_y - \frac{df_{y'}}{dx} = 0 \quad (7)$$

which ensures condition for extremum but does not tell whether it corresponds to minimum or maximum. To find these conditions, let us perform Taylor's expansion up to second order:

$$f(y' + \alpha\eta', y + \alpha\eta, x) = f(y', y, x) + \alpha\eta f_y + \alpha\eta' f_{y'} + \frac{(\alpha\eta)^2}{2!} f_{yy} + 2\frac{(\alpha^2\eta\eta')}{2!} f_{yy'} + \frac{(\alpha\eta')^2}{2!} f_{y'y'} \quad (8)$$

and therefore

$$\frac{\partial^2 I}{\partial \alpha^2}|_{\alpha=0} = \int_{x_1}^{x_2} [\eta^2 f_{yy} + 2\eta\eta' f_{yy'} + \eta'^2 f_{y'y'}]dx. \quad (9)$$

Now $\int_{x_1}^{x_2} \eta\eta' f_{yy'} dx = \int_{x_1}^{x_2} [\frac{d}{dx}(\frac{\eta^2}{2})] f_{yy'} dx$ which upon integration by parts gives $[f_{yy'}\frac{\eta^2}{2}]_{x_1}^{x_2} - \int_{x_1}^{x_2} [\frac{d}{dx}(f_{yy'})]\frac{\eta^2}{2} dx$ and since $\eta(x_1)$=0=$\eta(x_2)$, it is just $-\frac{1}{2}\int_{x_1}^{x_2} \eta^2 [\frac{d}{dx}(f_{yy'})]$. Thus

$$\frac{\partial^2 I}{\partial \alpha^2}|_{\alpha=0} = \int_{x_1}^{x_2} [\eta'^2 f_{y'y'} + \eta^2 \{f_{yy} - \frac{d}{dx}(f_{yy'})\}]dx. \tag{10}$$

We can choose $\eta(x)$ to be any arbitrary sawtooth function. Sawtooth function is a continuous function of x but its derivative $\eta'(x)$ is not, rather, η' is alternately $+m$ and $-m$ where m is a constant. But as η'^2 appears in the above equation, we have the advantage that $\eta'^2 = m^2$ which is always positive and remains fixed. Also, the sawtooth function η can be chosen as small as possible ($|\eta| < \epsilon$) while keeping η' same ($= \pm m$). Thus for $|\eta|$ arbitrarily small, it follows that $\frac{\partial^2 I}{\partial \alpha^2} > 0 \Rightarrow m^2 \int_{x_1}^{x_2} f_{y'y'} dx > 0$, and since m^2 is positive and the preceding arguments hold for any arbitrary fixed end points x_1 and x_2, the inequality

$$f_{y'y'} > 0 \tag{11}$$

represents the necessary condition that the extremisation by Eq.(7) provides minimum and *vice versa*. Therefore, whenever it is necessary to ascertain whether extremisation by Eq.(7) provides minimum or maximum, $f_{y'y'} > 0$ or $f_{y'y'} < 0$ should be checked.

2 Snell's law for waves and massive particles

Snell's law (also known as Snell-Descartes law and the law of refraction) is a formula used to describe the relationship between the angles of incidence and refraction, when referring to light or other waves passing through a boundary between two different isotropic media. It states that the ratio of the sines of the angles of incidence and refraction is equivalent to the ratio of phase velocities in the two media, or equivalent to the opposite ratio of the indices of refraction (with respect to vacuum) resulting in bending of a ray towards the normal (to the boundary separating the two media) for the medium in which velocity of light or other waves is less. The indices of refraction of the media, labeled n_1, n_2 etc. represent the factor by which a light (or other) ray's speed decreases when traveling through a refractive medium as opposed to its velocity in vacuum. Snell's law for waves can be readily proved by Fermat's principle of least time (taken by the light or other waves for traveling from one point to the other across a boundary) or derived from wave nature of light (or other waves) using Huygen's construction [2].

Let us now consider the case of a massive particle (such as neutron) which is incident (from vacuum) with kinetic energy E on a nucleus of radius R which offers a uniform potential $-V$ (where V is positive) to the incident particle. Although massive, if this particle is treated like wave and as its velocity inside the nucleus and vacuum are proportional to $\sqrt{E+V}$ (from energy conservation) and $\sqrt{E}$, respectively, it would result in bending of the ray (as described above) away from the normal (which is along the radius of the nucleus) inside the nucleus where velocity is more. This would lead to the existence of critical angle $sin^{-1}\sqrt{E/(E+V)}$ beyond which there is no transmission (even in an attractive potential of $-V$) and a refractive index n for the nuclear medium less than that of vacuum (which is unity) which are physically unacceptable. Thus, the case of a massive particle can not be treated as wave. For deriving Snell's law for massive particle one can use the principle of angular momentum conservation. If b (the impact parameter) and x are the lengths of the perpendiculars on the incident and deflected paths from the centre of the nucleus, the angular momentum conservation provides the relation $b\sqrt{E} = x\sqrt{E+V}$ where initial (incident) angular momentum is equated to the final angular momentum of the deflected particle. This immediately shows that $x < b$ implying that the refracted (deflected) particle bends towards the centre of the nucleus and the refractive index $n = \sqrt{E+V}/\sqrt{E}$ which is greater than unity. These physically correct results, which were also the intuitive results, are just the opposite of those if the particle were considered as a wave.

3 Summary and Conclusion

In this short note, the necessary conditions of minimum and maximum for extremisation by Euler-Lagrange equation are obtained and the Snell-Descartes law for massive particles is derived which is in contradiction to that for the light waves.

References

[1] H. Goldstein, Classical Mechanics, Chapter 2, Addison-Wesley Publishing Company Inc. (1950).

[2] F. A. Jenkins and H. E. White, Fundamentals of Optics, McGraw-Hill, New York, (1976).

Stern-Gerlach Experiments and Complex Numbers in Quantum Physics

S. Sivakumar

Materials Physics Division
Indira Gandhi Centre for Atomic Research
Kalpakkam 603 102 INDIA
siva@igcar.gov.in

(Submitted April 2012)

Abstract

It is often stated that complex numbers are essential in quantum theory. In this article, the need for complex numbers in quantum theory is motivated using the known results of tandem Stern-Gerlach experiments.
Keywords: Stern-Gerlach experiment, complex numbers, superposition

1 Introduction

Complex numbers are essential in quantum theory. In classical physics complex quantities are often introduced to aid in solving problems rather than as a necessity. That makes it mysterious for students about the role of complex numbers in quantum theory. In this pedagogical report, it is illustrated that the need for complex numbers in quantum theory can be made plausible after discussing the results of Stern-Gerlach (SG) experiment. This idea is presented in many texts, for instance, Sakurai[1] or Townsend[2]. Here, we wish to bring this to the notice of physics students and make a simplified presentation.

A SG apparatus is an arrangement to provide a spatially inhomogeneous magnetic field. The purpose of spatial inhomogeneity is to exert force on spins, which are like magnetic moments, so that spins of different orientations are spatially separated. The direction of maximum gradient (a measure of inhomogeneity) is the axis along which spatial separation occurs. If this direction is chosen to be the z-axis, the corresponding SG apparatus is said to be oriented along z-axis and it is denoted by SG_z. If the "spin" is indeed like a classical magnetic moment, then every possible orientation with respect to the orientation of the SG_z is possible and the output beam is expected to be continuously distributed along z direction in in space. However, experiments indicated that there were finite number of output streams. Particles in each of the stream is assigned a "spin" value. If there are two outputs, the particles in one of the beams are said to be in up-spin state and those in the other output are said to be in the down-spin state. Such particles are said to be "spin-half" particles. Electrons, protons, neutrons, singly ionized silver atoms are some examples of spin-half systems.

2 Tandem Stern-Gerlach Experiments

The need for introducing complex numbers is easily recognized by knowing the results of experiments using two SG apparatuses in tandem. Consider a beam of spin-half system, for example, singly ionized silver atoms, passing through a SG_z. The output of the apparatus will have two beams that are spatially separated. This indicates that the spin of an atom in the beam has two possible values. In quantum theory this is taken to mean that the required state space is two-dimensional. Associated with these two possible spin values are two quantum states, namely, $|z+\rangle$ and $|z-\rangle$, corresponding to up-spin and down-spin respectively. . An arbitrary spin state $|\psi_{in}\rangle$ is described by a superposition of the two states,

$$|\psi_{in}\rangle = r_1|z+\rangle + r_2|z-\rangle, \qquad (1)$$

where r_1 and r_2 are the superposition coefficients that satisfy $r_1^2 + r_2^2 = 1$. A short notation is used to present these facts. A SG apparatus oriented along the z-axis is denoted by Z enclosed in a box. The experimental fact that an arbitrary beam of spin-half systems will give rise to two output beams is represented by

$$|\psi_{in}\rangle \longrightarrow \boxed{\mathrm{Z}} \longrightarrow \{|z+\rangle, |z-\rangle\},$$

where the states corresponding to the two output beams are enclosed in curly brackets. The relative intensities of the output beams decide the magnitude of the superposition coefficients. Let us *assume that the superposition coefficients are real.* According to the Born's rule for statistical interpretation, the relative intensities of the beams corresponding to orthogonal states are the squares of the magnitudes of the respective superposition coefficients. In the case of SG_z experiment with two output beams of equal intensity, the input state is a superposition of the two output states:

$$|\psi_{in}\rangle = \frac{1}{\sqrt{2}}[|z+\rangle + |z-\rangle. \qquad (2)$$

If the input beam is in the state $|z+\rangle$, there is a single output beam corresponding to the output stream represented by $|z+\rangle$. In the short notation introduced earlier, tt is represented as

$$|z+\rangle \longrightarrow \boxed{\mathrm{Z}} \longrightarrow |z+\rangle.$$

That is, $|z-\rangle$ cannot be generated from $|z+\rangle$ using SG_z. Similarly, if the input state is $|z-\rangle$,

$$|z-\rangle \longrightarrow \boxed{\mathrm{Z}} \longrightarrow |z-\rangle,$$

implying that $|z+\rangle$ cannot be obtained from $|z-\rangle$. In simple terms, SG_z does not affect $|z+\rangle$ and $|z-\rangle$. Hence, they qualify as "eigenstates" of SG_z. More importantly, the fact that SG_z cannot generate $|z+\rangle$ from $|z-\rangle$ implies that the two states $|z+\rangle$ and $|z-\rangle$ are "orthogonal" to each other. In mathematical terms, orthogonality means the inner product between the two states is zero.

The choice of orientation of the SG apparatus is arbitrary. For instance, if the SG apparatus is oriented along x-direction, then an arbitrary input beam of spin- half particles results in two output beams, separated spatially along the x- direction. The respective states of the particles in the two beams are denoted by $|x+\rangle$ and $|x-\rangle$. As in the case of SG_z, the following are true:

$$|\psi_{in}\rangle \longrightarrow \boxed{\mathrm{X}} \longrightarrow \{|x+\rangle, |x-\rangle\},$$

$$|x+\rangle \longrightarrow \boxed{\mathrm{X}} \longrightarrow |x+\rangle,$$

and

$$|x-\rangle \longrightarrow \boxed{\mathrm{X}} \longrightarrow |x-\rangle.$$

And the conclusion is that the states $|x+\rangle$ and $|x-\rangle$ are orthogonal, eigenstates of SG_x. Similarly, for an experiment with SG_y ,

$$|\psi_{in}\rangle \longrightarrow \boxed{\mathrm{Y}} \longrightarrow \{|y+\rangle, |y-\rangle\},$$

$$|y+\rangle \longrightarrow \boxed{\mathrm{Y}} \longrightarrow |y+\rangle,$$

and

$$|y-\rangle \longrightarrow \boxed{\mathrm{Y}} \longrightarrow |y-\rangle.$$

As in the other cases, the states $|y+\rangle$ and $|y-\rangle$ are orthogonal, eigenstates corresponding to SG_y.

2.1 Experiment I

Are $|z+\rangle$ and $|z-\rangle$ unaffected by SG_x? To find out, one of the outputs of SG_z, say, the beam of particles corresponding to $|z+\rangle$, is used as input to SG_x. The experimental result is that there are two output beams of equal intensity. So, from $|z+\rangle$, both $|x+\rangle$ and $|x-\rangle$ emerge. Then the following assignment is possible:

$$|z+\rangle = \frac{1}{\sqrt{2}}[|x+\rangle + |x-\rangle], . \tag{3}$$

Once this choice is made for $|z+\rangle$, the requirement for orthogonality implies that

$$|z-\rangle = \frac{1}{\sqrt{2}}[|x+\rangle - |x-\rangle]. \tag{4}$$

These expressions are consistent with the requirement that $|z+\rangle$ and $|z-\rangle$ are orthogonal to each other. Note that the superposition coefficients are chosen to be real. It does not matter if the expressions for the states $|z+\rangle$ and $|z-\rangle$ are swapped.

2.2 Experiment II

Let one of the outputs of SG_z be sent through a SG_y. Like the previous case, two output beams of equal intensity emerge from the apparatus. Arguing as before, the results are

$$|z+\rangle = \frac{1}{\sqrt{2}}[|y+\rangle + |y-\rangle], \quad (5)$$

$$|z-\rangle = \frac{1}{\sqrt{2}}[|y+\rangle - |y-\rangle], \quad (6)$$

where the superposition coefficients have been assumed to be real. There is no inconsistency so far.

2.3 Experiment III

The last piece of information required is a relationship among the states $|x\pm\rangle$ and $|y\pm\rangle$. For this, one of the output beams of SG_x, for instance, the output corresponding to $|x+\rangle$, is fed as input to SG_y. Two output beams of equal intensity emerge. If the input is changed to $|x-\rangle$, there are two output beams of equal intensity. So, the results can be summarized as

$$|x+\rangle = \frac{1}{\sqrt{2}}[|y+\rangle + |y-\rangle], \quad (7)$$

$$|x-\rangle = \frac{1}{\sqrt{2}}[|y+\rangle - |y-\rangle], \quad (8)$$

assuming that the superposition coefficients are real.

3 Analysis of results

What can be inferred from the results of the three experiments described above? First of all, the conclusions from the Experiment III can be used to rewrite the results of the Experiment II. This yields

$$|z+\rangle = |x+\rangle, \quad (9)$$

$$|z-\rangle = |x-\rangle. \quad (10)$$

This is at variance with the results of the Experiment I which indicate that $|z+\rangle$ and $|z-\rangle$ are linear combinations of $|x+\rangle$ and $|x-\rangle$. Obviously, one of the assumptions used in expressing the results should be wrong. The crucial assumption made is that the input state is expressible as a linear combination of output states with *real* coefficients. Now, it needs to be argued that using complex coefficients yields consistent results. The requirements are that the two output beams are of equal intensity and the corresponding states orthogonal to each other. So, one possibility is to recast the results of Experiment III using complex coefficients to give

$$|x+\rangle = \frac{1}{2}[(1-i)|y+\rangle + (1+i)|y-\rangle], \quad (11)$$

$$|x-\rangle = \frac{1}{2}[(1+i)|y+\rangle + (1-i)|y-\rangle]. \quad (12)$$

where $i = \sqrt{-1}$. The definition of inner product between two states $|\psi_1\rangle = a|z+\rangle + b|z-\rangle$ and $|\psi_2\rangle = c|z+\rangle + d|z-\rangle$ is $\langle\psi_1|\psi_2\rangle = a^*c + b^*d$, where superposition coefficients a, b, c and d are complex numbers, and the superscript $*$ implies complex conjugation. With this definition of inner product, the orthogonality condition is satisfied. Further, the coefficients are of equal magnitude to account for the observation that the output beams are of equal intensity. This specific choice of superposition coefficients ensures that the results of the Experiments I and II need not be rewritten with complex coefficients, and it concurs with the convention adopted in quantum physics. Other choices such as

$$|x+\rangle = \frac{1}{\sqrt{2}}[|y+\rangle + i|y-\rangle], \tag{13}$$

$$|x-\rangle = \frac{1}{\sqrt{2}}[|y+\rangle - i|y-\rangle], \tag{14}$$

to express the results of Experiment III would require rewriting the results of the Experiment I and Experiment II using complex superposition coefficients.

4 Discussion

Complex numbers are essential in the Hilbert space formulation of quantum theory. Without invoking complex numbers, it is impossible to consistently explain the outcomes of some simple experiments performed with SG devices in tandem. Another important point to note is that the Schrodinger equation has not been used in the arguments presented here. Even though $\sqrt{-1}$ appears explicitly in the Schrodinger equation which governs dynamics in quantum physics, the requirement for complex numbers is not due to this particular rule of dynamics. It is the linear vector space structure that is crucial in necessitating complex numbers in quantum theory.

References

[1] J J Sakurai, Introduction to Modern Quantum Mechanics(Addison-Wesley, 1994, New York) p27.

[2] J. S. Townsend, A Modern Introduction to Quantum Mechanics (McGraw Hill, 1992, Singapore) p17.

Physics Through Problem Solving - XXIV
Poisson Brackets

Ahmed Sayeed

Department of Physics
University of Pune
Pune - 411007
sayeed@physics.unipune.ac.in

(Submitted September 8 2012)

Abstract

In this installment we shall do some problems with Poisson brackets. The Poisson brackets can be used to state the equation of motion (i.e., time dependence in the form of a differential equation) of any function of coordinates and momenta (i.e, 'a dynamical variable') in a very elegant manner which emphasizes the role played by the Hamiltonian function and the constants of motion. The problems are meant to demonstrate these aspects of Poisson brackets

Consider a system n degrees of freedom, whose phase space coordinates are $\mathbf{q} = \{q_1, q_2, \ldots, q_n\}$ and $\mathbf{p} = \{p_1, p_2, \ldots, p_n\}$. The Poisson bracket (PB for short) of two dynamical variable of this system, $u(\mathbf{p}, \mathbf{q}, t)$ and $v(\mathbf{p}, \mathbf{q}, t)$, is defined as

$$[u, v] = \sum_i \left(\frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i} \right) \qquad (1)$$

The range of summation is from $i = 1$ to $i = n$, and will be assumed in all the following expressions, unless otherwise mentioned. But the index of summation will be always mentioned (no summation convention used anywhere).

Here we summarize some basic properties of Poisson brackets which we will be using in solving the problems of this issue. In the following u, v and w are three dynamical variables and c is a constant.

$$[u,v] = -[v,u] \quad (2)$$
$$[u,u] = 0 \quad (3)$$
$$[cu,v] = [u,cv] = c\,[u,v] \quad (4)$$
$$[u+v,w] = [u,w]+[v,w] \quad (5)$$
$$[u,v+w] = [u,v]+[u,w] \quad (6)$$
$$[uv,w] = u[v,w]+[u,w]v \quad (7)$$
$$[u,vw] = v[u,w]+[u,v]w \quad (8)$$

Problem 1: Find the PBs $[L_x,L_y],[L_y,L_z]$ and $[L_z,L_x]$, where L_x,L_y,L_z are the Cartesian components of the angular momentum of a particle.

Solution: We have $L_x = yp_z - zp_y$, $L_y = zp_x - xp_z$, and $L_z = xp_y - yp_x$, where p_x,p_y,p_z are canonically conjugate momenta for the Cartesian coordinates x,y, and z respectively. For convenience, let us rename the variables as follows: $\{x,y,z\} \to \{x_1,x_2,x_3\}$, $\{p_x,p_y,p_z\} \to \{p_1,p_2,p_3\}$, and $\{L_x,L_y,L_z\} \to \{L_1,L_2,L_3\}$. With this notation we have $L_1 = x_2p_3 - x_3p_2$, $L_2 = x_3p_1 - x_1p_3$, $L_3 = x_1p_2 - x_2p_1$.

A compact and elegant method of carrying out the algebra to find these PBs is by using Levi-Civita symbol ϵ_{ijk} and the Kronecker delta symbol δ_{ij}. This you can find in some text books, for example in Rana & Joag [1]. Here we shall carry out the algebra without using these neat mathematical devices. In fact, we shall calculate only the first PB $[L_1,L_2]$, and the other two can be readily calculated in the same manner by the reader.

$$[L_1,L_2] = \sum_{i=1}^{3}\left(\frac{\partial L_1}{\partial x_i}\frac{\partial L_2}{\partial p_i} - \frac{\partial L_1}{\partial p_i}\frac{\partial L_2}{\partial q_i}\right)$$

(using eq. 1)

In the above sum the first two terms vanish as $\frac{\partial L_1}{\partial x_1} = \frac{\partial L_1}{\partial p_1} = \frac{\partial L_2}{\partial p_2} = \frac{\partial L_2}{\partial x_2} = 0$. Thus

$$\begin{aligned}[L_1,L_2] &= \left(\frac{\partial L_1}{\partial x_3}\frac{\partial L_2}{\partial p_3} - \frac{\partial L_1}{\partial p_3}\frac{\partial L_2}{\partial x_3}\right)\\ &= x_1p_2 - x_2p_1 = L_3\end{aligned}$$

In the same manner we get $[L_2,L_3] = L_1$ and $[L_3,L_1] = L_2$. Note the cyclical order of the indices. If the order is not cyclical we get negative signs, e.g., $[L_2,L_1] = -L_3$, from the anti-commutative property of PBs as stated in eq. 2. Also, by the property given by eq. 3 (which is actually a corollary of eq. 2) we have $[L_1,L_1] = [L_2,L_2] = [L_3,L_3] = 0$. We note that the quantum analogue of these brackets (commutator brackets) are given by $[L_1,L_2] = i\hbar L_3$ and so forth.

Problem 2: Using the Poisson theorem for PBs show that the angular momentum (about the centre of force) is a constant of motion for the motion of a particle under an inverse square law force.

Solution: Poisson theorem (also called Poisson's first theorem on PBs) states that for a dynamical variable $u(\mathbf{q},\mathbf{p},t)$

$$\frac{du}{dt} = [u,H] + \frac{\partial u}{\partial t} \quad (9)$$

This is actually the equation of motion of u. By definition u is a constant of motion if $\frac{du}{dt} = 0$.

The angular momentum **L** is a vector, and to show that it is conserved to we have to show that all the three components L_x, L_y and L_z are conserved. These components have no explicit time dependence, so the partial derivatives $\frac{\partial L_x}{\partial t} = \frac{\partial L_y}{\partial t} = \frac{\partial L_z}{\partial t} = 0$. Thus, from eq. 9 the three components are conserved if the three PBs $[L_x, H] = [L_y, H] = [L_z, H] = 0$. We shall prove one of them, i.e, $[L_z, H] = 0$, and the reader can easily prove the other two in the same manner. Once again for the components of position, momentum and angular momentum we shall use the notation used in problem 1.

We shall use Cartesian coordinates with the centre of force at the origin. The Hamiltonian is given by

$$\begin{aligned} H(x_1, x_2, x_3, p_1, p_2, p_3) &= \frac{1}{2m}\left(p_1^2 + p_2^2 + p_3^2\right) - \frac{k}{r} \\ &= \frac{1}{2m}\left(p_1^2 + p_2^2 + p_3^2\right) - \frac{k}{\sqrt{x_1^2 + x_2^2 + x_3^2}} \end{aligned} \tag{10}$$

where k is a constant, positive for attractive force and negative for repulsive force.

$$[L_3, H] = \sum_{i=1}^{3}\left(\frac{\partial L_3}{\partial x_i}\frac{\partial H}{\partial p_i} - \frac{\partial L_3}{\partial p_i}\frac{\partial H}{\partial x_i}\right) \tag{11}$$

We have the partial derivatives

$$\frac{\partial L_3}{\partial x_1} = p_2,\ \frac{\partial L_3}{\partial x_2} = -p_1,\ \frac{\partial L_3}{\partial x_3} = 0,$$

$$\frac{\partial L_3}{\partial p_1} = -x_2,\ \frac{\partial L_3}{\partial p_2} = x_1,\ \frac{\partial L_3}{\partial p_3} = 0,$$

$$\frac{\partial H}{\partial p_1} = \frac{p_1}{m},\ \frac{\partial H}{\partial p_2} = \frac{p_2}{m},\ \frac{\partial H}{\partial p_3} = \frac{p_3}{m},$$

$$\frac{\partial H}{\partial x_1} = \frac{kx_1}{r^{3/2}},\ \frac{\partial H}{\partial x_2} = \frac{kx_2}{r^{3/2}},\ \text{and}\frac{\partial H}{\partial x_3} = \frac{kx_3}{r^{3/2}}$$

Using these in eq. 11 we have $[L_3, H] = 0$.

Problem 3: Show that for a free particle moving in one dimension, the function $F = x - \frac{pt}{m}$ and $\frac{\partial F}{\partial t}$ are constants of motion. Here x, p, and m are position, momentum and mass of the particle. Do this by direct calculation of total time derivatives of F as well as $\frac{\partial F}{\partial t}$, and by using Poisson's first and second theorem about PBs.

Solution: Note that F is explicitly a function of time, but nevertheless it is a constant of motion. This is quite trivial to show by taking the total time derivative of F. We use that fact that for a free particle momentum p is a constant.

$$\begin{aligned} \frac{dF}{dt} &= \frac{dx}{dt} - \frac{p}{m} \\ &= \frac{p}{m} - \frac{p}{m} = 0. \end{aligned}$$

And

$$\frac{d}{dt}\left(\frac{\partial F}{\partial t}\right) = \frac{d}{dt}\left(-\frac{p}{m}\right) = 0$$

Now we use Poisson's first theorem stated above in the previous problem. Note that the Hamiltonian for the free particle is $H = \frac{p^2}{2m}$.

$$\begin{aligned}\frac{dF}{dt} &= [F,H] + \frac{\partial F}{\partial t} \\ &= \left(\frac{\partial F}{\partial x}\frac{\partial H}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial H}{\partial x}\right) - \frac{p}{m} \\ &= \left(\frac{p}{m} - 0\right) - \frac{p}{m} = 0\end{aligned}$$

And, as $\frac{\partial F}{\partial t} = -\frac{p}{m}$ is not an explicit function of time, if it is a constant of motion its PB with H must be zero, as we can see:

$$\begin{aligned}\left[\frac{\partial F}{\partial t}, H\right] &= \left[-\frac{p}{m}, \frac{p^2}{2m}\right] \\ &= -\frac{1}{2m^2}\left[p, p^2\right] \\ &= -\frac{1}{2m^2}\left(p\,[p,p] + [p,p]\,p\right) = 0\end{aligned}$$

In the above we have used PB properties given by eqs. 3, 4 and 8 for illustration, though here it is equally simple to take the partial derivatives. We can also illustrate Poisson's second theorem on PBs, which states that the PB of two constants of motion is itself a constant of motion. Now that we know F is a constant of motion, we can take its PB with H, which is also a constant of motion (because the particle is free). This PB we already evaluated above, i.e., $[F,H] = \frac{p}{m} = -\frac{\partial F}{\partial t}$. So by Poisson's second theorem $\frac{\partial F}{\partial t}$ must be a constant of motion, as we already verified.

This demonstrates one valuable application of Poisson's second theorem: If we have two constants of motion, we can take their PB to construct one more constant of motion, which might of interest. But it can also turn out some function of already known constants of motion, which can hardly be of any interest. Consider this example. Here we have two constants of motion, F and $\frac{\partial F}{\partial t}$, and their PB is

$$\begin{aligned}\left[F, \frac{\partial F}{\partial t}\right] &= \left[x - \frac{pt}{m}, -\frac{p}{m}\right] \\ &= \left[x, -\frac{p}{m}\right] + \left[-\frac{pt}{m}, -\frac{p}{m}\right] \\ &= -\frac{1}{m}\,[x,p] - \frac{t}{m^2}\,[p,p] = -\frac{1}{m}\end{aligned}$$

(using the properties of PBs given by eqs. 3, 4 and 8, and $[x,p] = 1$), which is obviously a constant of motion, and not a terribly interesting one, as all it means is that mass remains constant during the motion.

Problem 4: Consider the following functions of position q and momentum p of a one-dimensional harmonic oscillator (m is the mass and ω angular frequency) :

$$a = \sqrt{\frac{m\omega}{2}}\left(q + \frac{ip}{m\omega}\right) \tag{12}$$

and its complex conjugate

$$a^* = \sqrt{\frac{m\omega}{2}}\left(q - \frac{ip}{m\omega}\right) \tag{13}$$

Write down and solve the equations of motion for a and a^* in terms of PBs of these functions with the Hamiltonian, and from these solutions find the solutions $q(t)$ and $p(t)$.

Solution: The equations of motion for a and a^* are (from Poisson's first theorem discussed above)

$$\frac{da}{dt} = [a, H] \tag{14}$$

and

$$\frac{da^*}{dt} = [a^*, H] \tag{15}$$

Note that the partial derivatives $\frac{\partial a}{\partial t}$ and $\frac{\partial a^*}{\partial t}$ are absent in the above equations of motion as both are zero, because the functions a and a^* are not explicitly time-dependent. The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 \tag{16}$$

We can evaluate the the PBs $[a, H]$ and $[a^*, H]$ by using the general definition given in eq. 1 or by using the properties of PBs listed in eqs. 2 – 8. We use the latter method, because that will also allow the reader to compare the functions a and a^* with their quantum mechanical analogues – the lowering and raising operators a and $a^\dagger$. Consider first the product

$$\begin{aligned} aa^* &= \sqrt{\frac{m\omega}{2}}\left(q + \frac{ip}{m\omega}\right)\sqrt{\frac{m\omega}{2}}\left(q - \frac{ip}{m\omega}\right) \\ &= \frac{m\omega}{2}\left(q^2 + \frac{p^2}{m^2\omega^2}\right) \end{aligned}$$

Comparing this with eq.16 we immediately get

$$H = \omega a a^* \tag{17}$$

Let us also get

$$\begin{aligned} [a, a*] &= \frac{\partial a}{\partial q}\frac{\partial a^*}{\partial p} - \frac{\partial a}{\partial p}\frac{\partial a^*}{\partial q} \\ &= \frac{m\omega}{2}\left(-\frac{i}{m\omega}\right) - \frac{m\omega}{2}\left(\frac{i}{m\omega}\right) \\ &= -i \end{aligned} \tag{18}$$

Now

$$\begin{aligned} [a, H] &= [a, \omega a a^*] \quad \text{(using eq. 17)} \\ &= \omega\,[a, aa^*] \quad \text{(using eq. 4)} \\ &= \omega\,(a\,[a, a^*] + [a, a]\,a^*) \ \text{(using eq. 8)} \\ &= -i\omega a \end{aligned} \tag{19}$$

In the last line we used

$$[a, a] = 0 \text{ and } [a, a^*] = -i.$$

Thus we have the equation of motion for a (eq. 14)

$$\frac{da}{dt} = -i\omega a \tag{20}$$

which can be immediately integrated to give

$$a = a_0 e^{-i\omega t} \tag{21}$$

where a_0 is the constant of integration.

Similar calculations give us $[a^*, H] = i\omega a^*$ and using it in the equation motion of eq. 15, and integrating we get

$$a^* = a_0^* e^{i\omega t} \tag{22}$$

where a_0^* is the constant of integration.

Now to find the solutions $q(t)$ and $p(t)$, we solve eqs. 12 and 13 for q and p to get

$$q = \sqrt{\frac{1}{2m\omega}}(a + a*) \tag{23}$$

and

$$p = -i\sqrt{\frac{m\omega}{2}}(a + a*) \qquad (24)$$

Now using eqs. 21 and 22 in eqs. 23 and 24 we have

$$q = \sqrt{\frac{1}{2m\omega}}\left(a_0 e^{-i\omega t} + a_0^* e^{i\omega t}\right) \qquad (25)$$

$$p = -i\sqrt{\frac{m\omega}{2}}\left(a_0 e^{-i\omega t} - a_0^* e^{i\omega t}\right) \qquad (26)$$

References

[1] N. C. Rana and P. S. Joag *Classical Mechanics* (Tata McGraw-Hill, New Delhi (1991).

COMPOUND PENDULUM

S. R. PATHARE, R. D. LAHANE, S. S. SAWANT*, C. C. PATIL
Homi Bhabha Centre for Science Education (TIFR)
V. N. Purav Marg, Mankhurd. Mumbai 400 088.
shirish@hbcse.tifr.res.in

*Bhavan's College, Munshi Nagar, Andheri (W), Mumbai – 400 058.
(*Submitted* 08-08-2012)

Abstract

This is an edited version of the experiment set for the experimental examination conducted at the orientation cum selection camp held at Homi Bhabha Centre for Science Education (TIFR), Mumbai in May 2010. Generally, the compound pendulum studied in undergraduate laboratory is in the form of a uniform bar whose axis of oscillation is varied. In this experiment, a compound pendulum with a fixed axis of oscillation but with a movable mass is used to study the dependence of periodic time on the position of the movable mass and to determine the gravitational field strength.

1. Introduction

The compound pendulum provided for this experiment consists of a rod with a fixed knife-edge, which acts as the axis of oscillation of the oscillating pendulum somewhere along its length. A cylindrical body of mass m_1 is used which can be moved along the length of the rod. Another cylindrical body of mass m_2 is fixed at the lower end of the rod. A plastic washer is used to support the mass m_1 at various positions on the rod.

When the pendulum is suspended with its knife-edge on a rigid platform and set into oscillation, its periodic time of oscillation changes depending on the position of the movable mass. The experiment consists of studying the relationship of the distance of the movable mass from the axis of oscillation with the periodic time of the pendulum.

2. APPARATUS

1) A compound pendulum consisting of a rod with one mass attached at one of its ends, another mass capable of sliding along the rod and a knife edge to be fixed on the rod,
2) An Allen key,
3) A plastic washer for supporting movable mass,
4) An acrylic support with fixed glass slides on which the knife edge is to rest,
5) A G-clamp for clamping the acrylic support to the edge of the table,
6) A stopwatch,
7) A measuring tape,
8) Vernier calipers and
9) A micrometer screw gauge.

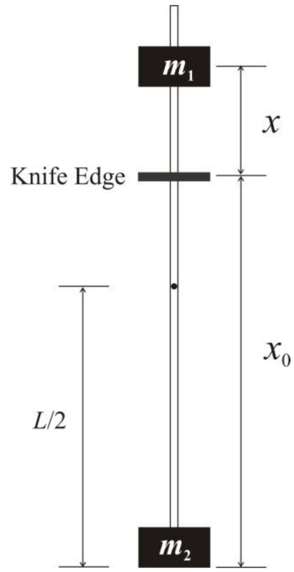


Figure 1.The compound pendulum

3. Description of apparatus:

Figure 2. A rod with a mass m_2 fixed at one end.

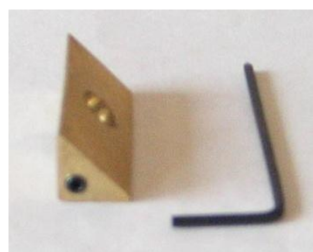


Figure 3. Knife Edge with Allen key

Figure 4. Acrylic support with glass slides fixed on it and the G-clamp

Figure 5. Stopwatch, measuring tape and vernier calipers and micrometer screw gauge

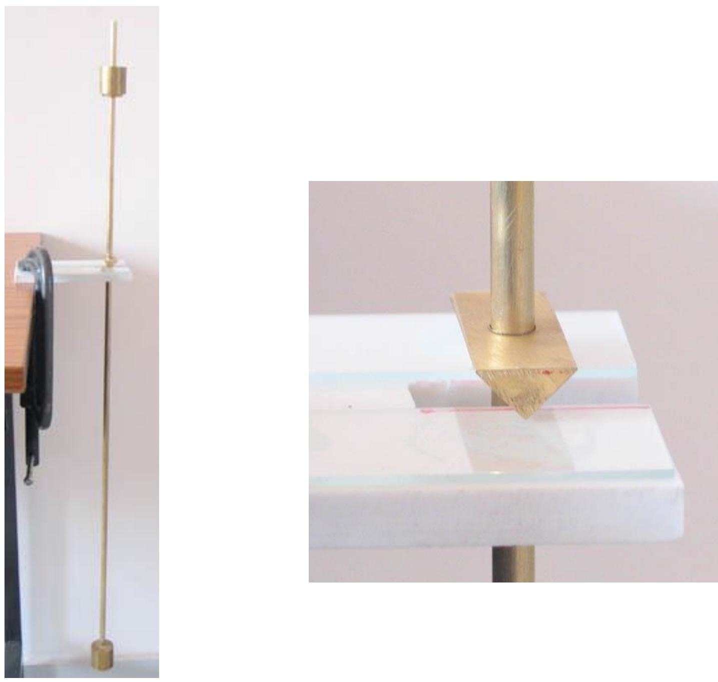


Figure 6. The complete setup

4. Theory

For small oscillations, the periodic time T of compound pendulum, with mass M and moment of inertia I about the axis of oscillation, is given to a good approximation by

$$T = \pi\sqrt{\frac{I}{Mgl}},$$

where, l is the distance between the axis of oscillation and the centre of mass of the pendulum.

5. Experiment

Given data:

Masses of

i) Rod: $m_r = 161 \pm 1$ g

ii) Knife-edge: $m_{ke} = 12 \pm 1$ g

iii) Bodies: $m_1 = m_2 = 99.5 \pm 0.5$ g

1. Express the distance l and the moment of inertia I about the axis of oscillation in terms of the distance x of the movable mass m_1 and other constants of the system.

Use the following symbols in your derivations.
Length of the rod: L.
Distance between the knife-edge and the end of the rod with fixed cylindrical body: x_0
Inner and outer radii of the cylindrical bodies: R_1 and R_2.
Length of the cylindrical bodies with masses m_1 and m_2: h.
Mass of the rod: m_r
Mass of the assembly of knife edge: m_{ke}.

2. Hence express the periodic time T of the pendulum as a function of x. If you have made any assumptions in neglecting any terms in the above derivations mention them with supporting arguments.

3. Make the necessary measurements of physical dimensions of the system forming the compound pendulum. You may use scales, vernier calipers and micrometer screw as required. Tabulate the measured values along with the uncertainties in measurements.

4. Suspend the pendulum from the rigid support and determine its periodic time for different positions of the movable mass by moving it from the top of the rod to the fixed mass at the bottom in suitable steps. [For moving the mass below the knife edge remove the knife edge using the Allen key and after shifting the mass below its position fix it again.] Tabulate your results.

5. Sketch graphs (rough sketches on the plain answer sheet) to show how I and l vary as the mass m_1 is shifted from one end to the other. Plot T versus x and explain the significance of the minimum T in this graph.

6. Reorganize the terms in the equation of T as a function of x and plot a linear graph from which g can be obtained. Determine the slope of the graph, calculate g and estimate the uncertainty in the obtained value.

7. Obtain from the graph the value of T at $x = 0$. Determine the value of g using the formula for T with $x = 0$.

8. If the movable mass is kept at the top end of the pendulum and you are allowed to move the axis of suspension, will it be possible to make T infinite? Explain the conditions under which this is possible. Will it be possible to achieve the condition experimentally? Substantiate your answer with reasons, if necessary.

6. TYPICAL OBSERVATIONS AND CALCULATIONS

1)

<u>Equation for ***I***</u>:

$$I = \frac{m_r}{12}L^2 + m_r\left(x_0 - \frac{L}{2}\right)^2 + \frac{m_1 h^2}{12} + \frac{m_1}{4}\left(R_1^2 + R_2^2\right) + m_1 x^2 + \frac{m_2 h^2}{12} + \frac{m_2}{4}\left(R_1^2 + R_2^2\right)$$
$$+ m_2\left(x_0 - \frac{h}{2}\right)^2 + I_{ke} \qquad (1)$$

<u>Equation for *l*:</u>

$$Ml = m_r\left(x_0 - \frac{L}{2}\right) + m_2\left(x_0 - \frac{h}{2}\right) + m_1 x - m_{ke} \times x_{ke}$$

$$\therefore l = \frac{1}{M}\left[m_r\left(x_0 - \frac{L}{2}\right) + m_2\left(x_0 - \frac{h}{2}\right) + m_1 x - m_{ke} \times x_{ke}\right] \qquad (2)$$

Here, $M = m_r + m_1 + m_2 + m_{ke}$

We can neglect the terms I_{ke} in equation (1) and $m_{\text{ke}}\, x_{\text{ke}}$ in equation (2) because they would be very small.

2.

$$T = 2\pi\sqrt{\frac{I}{Mgl}}$$

$$T^2 l = 4\pi^2 \frac{I}{Mg}$$

Reorganizing the terms in equation (1) we can write

$$I = m_1 x^2 + \left[m_r\left(\frac{L^2}{12} + \left(x_0 - \frac{L}{2}\right)^2\right) + \frac{(m_1 + m_2)h^2}{12} + \frac{(m_1 + m_2)}{4}\left(R_1^2 + R_2^2\right) + m_2\left(x_0 - \frac{h}{2}\right)^2\right]$$

The terms in the bracket are constant. Representing the constant by A,

$$I = m_1 x^2 + A$$

We can reorganize the terms in equation (2) as

$$l = \frac{m_1}{M}x + \frac{1}{M}\left[m_r\left(x_0 - \frac{L}{2}\right) + m_2\left(x_0 - \frac{h}{2}\right)\right]$$

Again, representing the terms in bracket by a constant B,

$$l = \frac{m_1}{M}x + B$$

$$\therefore T^2\left(B + \frac{m_1}{M}x\right) = \frac{4\pi^2}{Mg}\left(m_1 x^2 + A\right)$$

3.

Measurements of physical dimensions of the system:

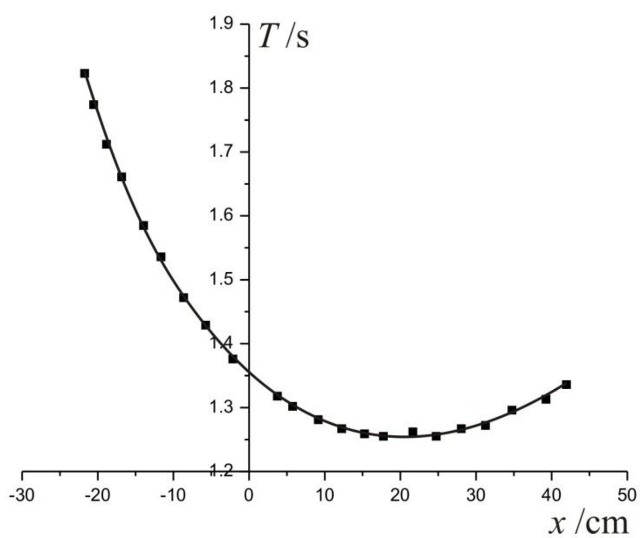
Quantity	Value	Uncertainty
L	69.0 cm	0.1 cm
h	2.500 cm	0.002 cm
x_0	45.9 cm	0.1 cm
R_1	1.265 cm	0.001 cm
R_2	0.300 cm	0.001 cm
Diameter of the rod	0.588 cm	0.002 cm

4.

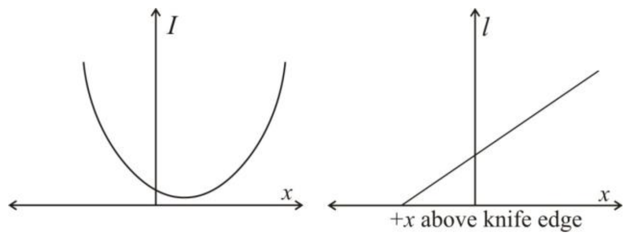
Obs. No.	$(x \pm \Delta x)$/cm	$x + \frac{h}{2}$ /cm	Time for 20 oscillations				T /s
			t_1 /s	t_2 /s	t_3 /s	Mean t /s	
1	-21.75 ± 0.10	-23.0	36.44	36.56	36.40	36.467	1.823 ± 0.008
2	-20.55 ± 0.10	-21.8	35.50	35.56	35.38	35.480	1.774 ± 0.009
3	-18.85 ± 0.10	-20.1	34.22	34.22	34.28	34.240	1.712 ± 0.003
4	-16.85 ± 0.10	-18.1	33.22	33.19	33.21	33.207	1.661 ± 0.002
5	-13.95 ± 0.10	-15.2	31.72	31.75	31.62	31.697	1.585 ± 0.006
6	-11.65 ± 0.10	-12.9	30.69	30.78	30.68	30.717	1.536 ± 0.005
7	-8.65 ± 0.10	-9.9	29.41	29.47	29.43	29.437	1.472 ± 0.003
8	-5.75 ± 0.10	-7.0	28.62	28.47	28.62	28.570	1.429 ± 0.008
9	-2.15 ± 0.10	-3.4	27.53	27.53	27.47	27.510	1.376 ± 0.003
10	3.75 ± 0.10	5.0	26.32	26.47	26.25	26.347	1.318 ± 0.011
11	5.75 ± 0.10	7.0	26.06	26.03	26.00	26.030	1.302 ± 0.003
12	9.15 ± 0.10	10.4	25.62	25.62	25.59	25.610	1.281 ± 0.002
13	12.25 ± 0.10	13.5	25.37	25.37	25.28	25.340	1.267 ± 0.004
14	15.25 ± 0.10	16.5	25.16	25.25	25.16	25.190	1.259 ± 0.004
15	17.75 ± 0.10	19.0	25.00	25.07	25.19	25.087	1.255 ± 0.010
16	21.65 ± 0.10	22.9	25.12	25.28	25.28	25.227	1.262 ± 0.008
17	24.75 ± 0.10	26.0	25.13	25.06	25.06	25.083	1.255 ± 0.004
18	28.05 ± 0.10	29.3	25.37	25.28	25.37	25.340	1.267 ± 0.004
19	31.25 ± 0.10	32.5	25.41	25.50	25.40	25.437	1.272 ± 0.005
20	34.75 ± 0.10	36.0	26.00	25.84	25.89	25.910	1.296 ± 0.008
21	39.25 ± 0.10	40.5	26.28	26.25	26.25	26.260	1.313 ± 0.002
22	41.95 ± 0.10	43.2	26.69	26.68	26.78	26.717	1.336 ± 0.005

5.

a. Plot *T* versus *x*.



b. Rough Sketches:



6.

$$\therefore T^2\left(\frac{m_1}{M}x+B\right)=\frac{4\pi^2}{Mg}\left(m_1x^2+A\right)$$

Plot a graph of $\therefore T^2\left(\frac{m_1}{M}x+B\right)$ versus x^2.

Here,

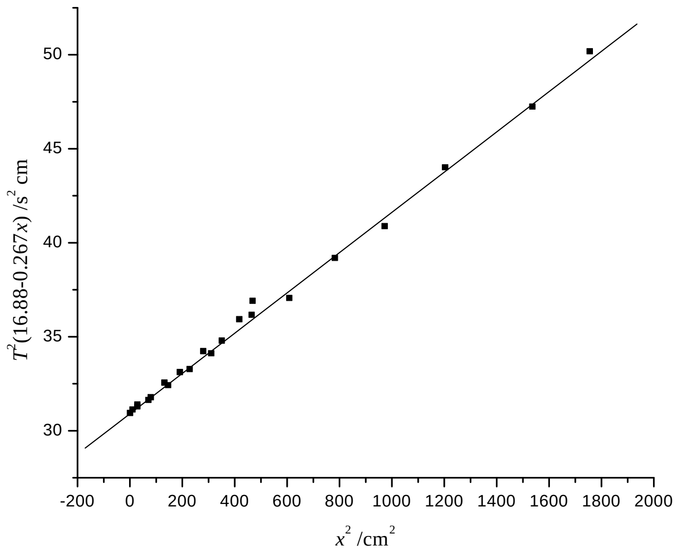
$$B=\frac{1}{M}\left[m_r\left(x_0-\frac{L}{2}\right)+m_2\left(x_0-\frac{h}{2}\right)\right]=\frac{1}{372}\left[161\left(45.9-\frac{69}{2}\right)+99.5\left(45.9-\frac{2.5}{2}\right)\right]$$

$$=\frac{1}{372}\left[1835.4+4442.7\right]=16.88$$

Hence, plot a graph of $T^2\left(16.88+0.267x\right)$versus x^2.

	$T^2(16.88+0.267x)$/cm s^2	x^2/cm^2
1	36.84	473
2	35.86	422
3	34.72	355
4	34.16	284
5	33.05	195
6	32.49	136
7	31.57	75
8	31.33	33
9	30.87	5
10	31.06	14
11	31.22	33
12	31.71	84
13	32.35	150
14	33.21	233
15	34.05	315
16	36.09	469
17	36.99	613
18	39.12	787
19	40.81	977
20	43.94	1208
21	47.17	1541
22	50.12	1760

Graph:



Slope = 0.0107

$$\text{Slope} = \frac{4\pi^2 m_1}{Mg} = 0.0107$$

$$g = \frac{4\pi^2 \times 99.5}{372 \times 0.0107} = 985.861 \ \text{dynes/g}$$

Uncertainty in g:

$$\frac{\Delta g}{g} = \sqrt{\left(\frac{\Delta m_1 / \sqrt{3}}{m_1}\right)^2 + \left(\frac{\Delta M / \sqrt{3}}{M}\right)^2 + \left(\frac{1}{\sqrt{3}}\frac{\Delta slope}{slope}\right)^2}$$

$$= \sqrt{\left(\frac{0.5 / \sqrt{3}}{99.5}\right)^2 + \left(\frac{1.6 / \sqrt{3}}{372}\right)^2 + \left(\frac{1}{\sqrt{3}} \times 0.0128\right)^2} = 0.00832$$

The expanded uncertainty

$$\Delta g = 0.00832 \times 985.861 \times 2 = 16.4 \cong 17 \ \text{dynes/g}$$

$$g = 986 \pm 17 \ \text{dynes/g}$$

7.

From the graph of T versus x:

$$T \text{ (at } x = 0) = 1.355 \text{ s}$$

$$\therefore T^2\left(B - \frac{m_1}{M}x\right) = \frac{4\pi^2}{Mg}\left(m_1 x^2 + A\right)$$

At $x = 0$

$$\therefore BT^2 = \frac{4\pi^2 A}{Mg}$$

Here

$$A = m_r\left(\frac{L^2}{12} + \left(x_0 - \frac{L}{2}\right)^2\right) + \frac{m_1 h^2}{6} + \frac{m_1}{2}\left(R_1^2 + R_2^2\right) + m_2\left(x_0 - \frac{h}{2}\right)^2$$

$$= 161\times\left(\frac{69^2}{12} + \left(45.9 - \frac{69}{2}\right)^2\right) + \frac{99.5\times 2.5^2}{6} + \frac{99.5}{2}\left(1.27^2 + 0.30^2\right) + 99.5\times\left(45.9 - \frac{2.5}{2}\right)^2$$

$$= 84800 + 104 + 85 + 198365 = 283354 \text{ g}\cdot\text{cm}^2$$

and

$$B = \frac{1}{M}\left[m_r\left(x_0 - \frac{L}{2}\right) + m_2\left(x_0 - \frac{h}{2}\right)\right] = \frac{1}{372}\left[161\left(45.9 - \frac{69}{2}\right) + 99.5\left(45.9 - \frac{2.5}{2}\right)\right]$$

$$= \frac{1}{372}\left[1835.4 + 4442.7\right] = 16.88$$

$$\therefore 16.88\times 1.355^2 = \frac{4\pi^2\times 283354}{372\,g}$$

$$g = 969 \text{ dynes/g or } 9.69 \text{ N/kg}$$

8.

If l is zero, the period will be infinite. This condition can be satisfied if the knife edge is moved and placed at the center of mass with both masses at the two ends of the rod. Another way can be to move the mass m_1 such that the centre of mass of the system coincides with the position of the knife edge.

In either of the case, when the pendulum is displaced from its equilibrium position, it will not return back to the equilibrium position as T is infinite. But the condition of unstable equilibrium will make it unrealizable experimentally.

7. Discussion

In the undergraduate laboratories, bar pendulum is a regular experiment. In that experiment, distance between the knife edge and centre of mass is varied in definite steps.

This experiment explores another way to study the compound pendulum by varying the position of centre of mass by shifting mass m_1 rather than the point of suspension.

The linearization technique in this experiment (to plot the suitable graph) requires a rearrangement of variables which itself is a skill to be developed by students.

8. Acknowledgements

We would like to thank Prof. D. A. Desai for his valuable guidance in the development of this experiment. We also wish to thank the students who helped us in standardizing the apparatus.

Quantum view of Mass

R. Ramachandran

Institute of Mathematical Sciences*, Taramani
Chennai, TN 600113, India.
rr@imsc.res.in, rr_1940@yahoo.co.in.unipune.ac.in

(Submitted 25-08-2012, revised 08-09-2012)

Abstract

The classical view of mass is that it quantifies the amount of substance and is a kinematical parameter. All matter has an attribute of mass and is a conserved quantity in any interaction. With the advent of Special Relativity, mass became no longer a conserved quantity, since Energy and Momenta had the status of conserved variables. Nevertheless, $\sqrt{(E^2 - \mathbf{p}^2c^2)} = mc^2$ gives a Poincare invariant measure that can be associated as the mass, an useful attribute of the body or system. In the quantum regime mass becomes truly dynamical. Higgs field is said to provide mass for all species of elementary constituents as widely popularized by the media in connection with the recent (most likely) discovery of Higgs boson at CERN. However, we emphasize that the most abundant component of matter Nucleons - derive their mass largely (95%) as a consequence of quantum effects of (color gluonic QCD) radiation. Further, interestingly this arises out of literally nothing, save the QCD scale, determined experimentally through a self consistent perturbative analysis of nucleon structure, as the sole input.

1 Introduction

Higgs particle discovery [1],[2],[3] has received much coverage and a perception that the so called God Particle is responsible to give

*Affiliated as retired Faculty; Communication address: Flat 12, Khagol Society, Panchvati, Pashan, Pune 411 008, India masses to all particles that make up the fundamental building blocks is prevalent. The object of this note to give a more appropriate perspective and provide non experts, particularly Physics Teachers and Students, a deeper view on what constitutes mass and what has been understood so far[1].

In classical physics, mass is a kinematical attribute of all matter. It is a measure of the quantity of matter and is perceived through two laws, both attributed to Newton. Force causes matter to accelerate and the proportionality constant is termed its inertial mass. Matter is also source of gravitational field it carries with it. This field falls off in intensity as Inverse Square of the distance from the source and the proportionate constant here is its gravitational mass. Galileo's famous experiment (at the leaning tower of Pisa?) and many modern equivalents demonstrate identity between the two definitions of mass and this implies a notion of universality of all bodies under gravitation. In classical regime the mass is a passive kinematical parameter and is conserved in any interaction. As we move to relativistic regime, we find that it is not mass that is conserved, but the Momentum (vector) $\mathbf{p}$ and Energy E. There is, however an invariant mass for every body or system which is given by $\sqrt{(E^2 - \mathbf{p}^2 c^2}/c^2$. Even this (Poincare invariant) mass is not conserved in any interaction, since mass of the system can be released as energy, heralding the celebrated relationship $E = mc^2$. In the terminology of Nuclear Physics, the mass defect shows up as the binding energy of nucleons in nuclei. Lighter nuclei such as Hydrogen, Helium and Carbon fuse to form tighter bound nucleus releasing the difference in mass as thermonuclear energies in a fusion reaction; and heavy nuclei, such as Uranium and Plutonium can be induced to undergo fission into medium heavy nuclei, releasing useful atomic energy, making in the process the iron region nuclei with highest binding energy per nucleon.

In quantum regime, we see that mass, whatever it may be, is dynamically generated. The notion of mass defect is an indication that the mass of a system, say an atom or nucleus is made up by a combination of the intrinsic mass of the constituents suitably dressed by interactions. The system may have a higher or lower mass than the sum total of constituents, making it either a resonant state or a bound state. For example, the energy spectrum of an atom is a consequence of the interaction of the constituents. Electromagnetic interaction between positively charged nucleus and negatively charged electrons results allowed energy levels in the atomic spectrum. To begin with, we have Schroedinger equation in Quantum mechanical description of an atom, say Hydrogen, give observed values of its spectra. This can be further improved and made fully relativistic in the language of Quantum Field Theory. Relevant field theory to deal with atomic (and molecular) spectra is Quantum Electrodynamics (QED), which comes endowed with Gauge symmetry. Gauge symmetry is a formal way of implementing a notion that the

[1] An excellent review is provided by F Wilczek [4]; supplement to 2011 Solvay conference, amplifies the content of this note

Electric and Magnetic field, that enters in the Lorentz Force law $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ is expressible in terms of scalar and vector potentials ($\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$; $\mathbf{B} = \nabla \times \mathbf{A}$). There is a freedom in the choice, since $\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla\chi(\mathbf{x}, t)$ and $\phi \to \phi' - \frac{\partial \chi(\mathbf{x},t)}{\partial t}$ leaves $\mathbf{E}$ and $\mathbf{B}$ invariant. Gauge theories formulated in terms of potential functions instead of field functions, necessary in quantum description (since Aharonov - Bohm [5] effect shows that the quantum electrons passing through magnetic field free region indeed detects change in the interference pattern, making potentials more fundamental than field strengths), comes endowed with a symmetry so generated. Since there is one function that characterises this symmetry, mathematically this is represented by a unitary unimodular group $U(1)$. We are able to achieve highly precise computations of the Energy levels (or masses) and transition rates, thanks to very reliable perturbation techniques, developed in the later half of the last century.

Proceeding further, nucleons in the nuclei are bound together by strong nuclear forces and the nucleons are indeed made up of quarks, bound by interactions mediated by gluons. Quarks and Leptons (electrons and the siblings) are the building blocks of all matter in the Standard Model. Like QED, the Standard Model is also a Gauge field theory with an underlying local [2] symmetry described by a symmetry group. While QED with symmetry group $U(1)$, [using relativistic four dimensional potentials A_μ that combines $\mathbf{A}$ and ϕ, a one parameter change $A_\mu \to A'_\mu = A_\mu - ie\partial_\mu \chi$ leaves the field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ unchanged] admits one (well known electric) charge and one gauge field whose quanta are photons, the Standard model has underlying symmetry group as $SU(3) \times SU(2) \times U(1)$ admits 3 coupling 'constants', 8+3+1 parameter symmetry transformations and has force fields as due to gauge bosons (8 gluons, 3 weak bosons and photon). They govern strong, weak and electromagnetic interactions of all basic constituents. An important difference is that while $U(1)$ of QED is commuting symmetry group (where two symmetry operations, one following the other in either order gives same result), the gauge group of Standard Model has non commuting components $SU(2)$ for weak interactions and $SU(3)$ of QCD for strong interactions. Here symmetry operations are represented by unitary unimodular 2×2 and 3×3 matrices. These are similar in character to rotations in space, which we know, in three or more dimensions, to be non-commuting. We may refer them as operations in some internal weak isospin and color space.

2 Massless Start

Theoretical description of basic interaction employs the tools of Relativistic Quantum Field Theory in the form of Gauge Theory. There are three pillars on which it stands and each of which needs, to begin with, massless

[2]The term 'local' implies the symmetry transformation parametrised by $\chi(\mathbf{x}, t)$ is a spacetime dependent function. If χ is a constant value independent of space and time, we will have a 'global' symmetry, such as flavour Isospin

fields as basic input.

2.1 Scale Invariance and Renormalization

We need our theories to be so that all observables yield finite values. It is necessary to prevent divergences, if any, from appearing in any measureable variable. This technical requirement is achieved by the process of Regularisation and Renormalization and this program is successful on account of the theory possessing scale invariance. In a scale invariant theory it is possible to promote the coupling constants, such as the 'fine structure constant' α here, into scale dependent parameters. The constant $\alpha = e^2/4\pi\hbar c$ becomes $\alpha(Q^2)$ and it measures the coupling strength or charge at different scales.

It is said that the 'Vacuum polarization' causes the bare charge to be screened, making charge depend on the scale of the probe used. A simple way to understand renormalisation is to note that in the quantum regime, 'vacuum' is anything but simple, since it can be thought of as all types of particle and antiparticle pairs to be continually created and annihilated perpetually, making it a polarisable medium or an effective dielectric. Just as effective charge in a dielectric medium gets reduced by the dielectric constant of the medium, a negatively charged electron with bare charge e_0 will polarise the nearby region of the 'vacuum' and consequently the measured charge will be the screened value. $e(Q^2)$ will be the effective charge when we approach it with a probe that causes a momentum transfer Q^2. Larger the value of Q^2, closer we approach it and lesser the screening. The value $e = 1.6 \times 10^{-19}$ Coulomb or $\alpha = 1/137$ is indeed the long range Thomson limit, when $Q^2 = 0$. For all this to make sense, the theory must possess an intrinsic scale invariance. A closely related symmetry is the angle preserving conformal invariance. Since, angle is ratio between two lengths, it does not change under scale transformation that varies all lengths in the same way. It is often convenient to use a set of units, such that $\hbar = 1 = c$ and in such units dimension of mass is just the inverse of the dimension of length. Recall the Compton wavelength λ associated with mass m is given by $\lambda = \hbar/mc$. If the theory has mass parameter, it possesses an intrinsic length; obviously such a theory can not be scale invariant. Thus basic ingredients in a renormalizable theory have necessarily to be massless. Presence of mass will imply scale violation.

`running_coupling.jpg`

Figure 1: $1/\alpha(Q^2)$ vs *log* Q^2

2.2 Chiral Fermions

Among the basic interactions are weak interactions, responsible for radioactivity. As early as 1957, we knew these to be parity violating. In order that this is so, we need to differentiate between the left handed and right handed states of fermions. In the Standard Model the left helicity states of quarks and leptons are doublets (in weak isospin, not to be confused with the more widely known flavor isospin), $Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$, $\psi_L = \begin{pmatrix} \nu \\ e_L^- \end{pmatrix}$ and the right helicity states of fermions u_R, d_R and e_R are singlets. Neutrino occurs in left helicity state only. If we view the fermion from a frame that moves faster than the particle (which is possible only if the particle has mass and travels with speed $v < c$) we will find in that frame, the helicity of the particle is reversed. Thus a chiral fermion, which is forced on us by the parity violation, is not compatible if the fermion has a mass. We require our fermions to be massless in order that they are viewed as chiral fermions. Fermionic matter consists of three copies (or three generations) of the above set, which is again forced on us as a need to accommodate a baryon asymmetric universe, which is populated mostly with nucleons with negligible fraction (10^{-10}) of anti-nucleons. That is another story. Neutrinoes also seem to mix and oscillate, which is possible when they have a tiny mass. That is yet another story.

2.3 Gauge interaction and massless bosons

The Standard Model describes interactions governed by the gauge theory with symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The subscripts C, L, Y refer to color (strong/chromodynamic), weak Left helicity isospin and a weak hypercharge Y respectively. Correspondingly there are gauge bosons; octet of gluons, electroweak bosons W^i, i =1, 2, 3 and W^0. $W^3$and W^0 combine to form two orthogonal combinations, of which one is the familiar electromagnetic photon γ, call it A and the other neutral weak boson Z. Together with $W^{1 \pm i2} = W^{\pm}$, we have weak intermediate bosons that mediate both neutral and charge changing weak interactions.

The Lagrangian density of the Standard model is given as:

$$
\begin{aligned}
\mathcal{L} \;=\; & -\frac{1}{4} G^a_{\mu\nu} G_a^{\mu\nu} - \frac{1}{4} W^i_{\mu\nu} W_i^{\mu\nu} - \frac{1}{4} W^0_{\mu\nu} W_0^{\mu\nu} \\
& - \overline{Q_L} \gamma^\mu (\partial_\mu - i g_3 \lambda_a G^a_\mu - i g_2 \frac{\tau_i}{2} W^i_\mu - \frac{1}{6} i g_1 W^0_\mu) Q_L \\
& + \overline{u_R} \gamma^\mu (\partial_\mu - \lambda_a G^a_\mu - \frac{2}{3} i g_1 W^0_\mu) u_R + \overline{d_R} \gamma^\mu (\partial_\mu - \lambda_a G^a_\mu + \frac{1}{3} i g_1 W^0_\mu) d_R \\
& + \overline{\psi_L} \gamma^\mu (\partial_\mu - i g_2 \frac{\tau_i}{2} W^i_\mu + \frac{1}{2} i g_1 W^0_\mu) \psi_L + \overline{e_R} (\partial_\mu + i g_1 W^0_\mu) e_R
\end{aligned}
$$

where $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + f_{abc} G^b_\mu G^c_\nu$, with a, b, c taking values $1, 2..8$, f_{abc} being the structure constants of $SU(3)$; $W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + \epsilon_{ijk} W^j_\mu W^k_\nu$, here i, j, k assume values $1, 2$ and 3, ϵ_{ijk} are structure constants for $SU(2)$; and $W^0_{\mu\nu} = \partial_\mu W^0_\nu - \partial_\nu W^0_\mu$.

G^a_μ is the octet vector field of of Gluons of QCD. $W^\pm_\mu = W^1_\mu \pm i W^2_\mu$ are the charged intermediate vector bosons that couple to charge changing weak currents, responsible for radioactivity (β decays); and the vector bosons W^3_μ and W^0_μ combine to become the conventional photon field $A_\mu (= \frac{g_1 W^3_\mu + g_2 W^0_\mu}{\sqrt{g_1^2 + g_2^2}})$ and the orthogonal neutral gauge boson Z_μ that is responsible for neutral weak current interactions. The coefficients of g_1 in the equation above reflect the weak hypercharge Y of the fermion field in that term. The coupling parameters g_i, as discussed in the preceding section on account of the renormalization process, get promoted into scale dependent functions $g_i(Q^2)$, where Q^2 is the square of the momentum transfer used to probe and their evolution as a function of Q^2 depends on what is known as the beta function $[\partial g / \partial log Q^2 = \beta(g)]$, of the respective symmetry group. A characteristic feature of these functions is that they make g_3 and g_2 logarithmically decrease as Q^2 increases, [while in contrast we have logrithmically increasing property for g_1] reflecting thus the anti-screening of the non-abelian charges. At extremely short distances, which need high momentum transfers and hence high energies to probe, the coupling is asymptotically vanishing. See the sketch in fig 1. Quarks color interactions are then small, amenable to perturbation treatment. Quark interactions are said to enjoy asymptotic freedom [6], [7] Deep inelastic scattering (high energy, high momentum transfer) by e or μ off nucleon targets revealed that quarks inside the nucleons can be regarded as free and non interacting!

Notice that there is no term quadratic in the gauge fields, such as $G^a_\mu G^\mu_a$, $W^i_\mu W^\mu_i$or $W^0_\mu W^\mu_0$, signifying that gauges bosons are like massless photons. There is no way to introduce a gauge preserving mass term. However, if the intermediate vector boson is massless, this will make the weak radioactivity a long ranged effect like electromagnetism! The mechanism to give masses to gauge bosons (without ruining the gauge symmetry), so that weak interactions remain short ranged is the celebrated Higgs mechanism. It achieves two outcomes. It makes the symmetry hidden (also referred to as spontaneously broken) in a way that the solution of the theory reflects a lesser symmetry (in our case $SU(3)_C \times U(1)_{em}$ than that of the underlying Lagrangian. The gauge bosons associated with the so called hidden symmetries, for us $W^\pm$ and Z, acquire mass. Further through the coupling the Higgs field has with all matter fermions, it also generates their masses. The minimum Higgs scheme calls for a new complex (weak isospin $\frac{1}{2}$) doublet scalar field $\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$, which is a color singlet and carries one unit of weak hypercharge Y. Together with its hermitean conjugate, we now have 4 scalar fields added through the Higgs phenomenon, with Lagrangian density (note wrong sign of

$\Phi^\dagger\Phi$ mass term)as

$$\mathcal{L}_{Higgs} = -[D_\mu\Phi^\dagger D^\mu\Phi - \mu^2(\Phi^\dagger\Phi) + \lambda(\Phi^\dagger\Phi)^2]$$

where the covariant derivative term

$$D_\mu\Phi = (\partial_\mu - ig_2\frac{\tau_i}{2}W_\mu^i - \frac{1}{2}ig_1W_\mu^0)\Phi$$

The shape of $V(\Phi) = -\mu^2(\Phi^\dagger\Phi) + \lambda(\Phi^\dagger\Phi)^2$,

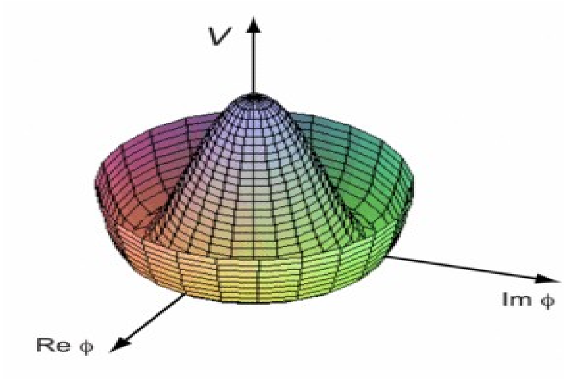


Figure 2: $V(\Phi)$ vs Φ illustrating degerate vacuum states

resembles a mexican hat as shown in the figure 2. On extremizing $V(\Phi)$, we find $\langle\Phi\rangle = 0$ as an unstable maxima and a degenerate set of minima, each of which can be the vacuum state, all characterized by a nonzero value of the Higgs field. We may choose the vacuum state to be given by

$$\langle\Phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix}, \text{where } v = \frac{\mu}{\sqrt{\lambda}}.$$

It is easily verified that the unbroken generator is $I_L^3 + Y/2$, that links the set of vacuum states and we associate it with the electromagnetism. Presence of non-vanishing v gives masses to $W^\pm$ and Z that are indeed observed at CERN with mass values 80 GeV and 91 GeV respectively. By a field redefinition we can demonstrate that three of the four Higgs fields metamorphose into the longitudinal modes of $W^\pm$ and Z bosons (now that these gauge bosons are massive they should have all three spin polarizations as against there being only two transverse polarizations for (massless) radiation), leaving one surviving mode, the recently discovered [1],[2] Higgs boson at 125 Gev.

Higgs coupling with the fermions (Yukawa interaction) provides masses for all fermions, such as quarks and leptons, the value of the mass being proportionate to the coupling parameter. Yukawa terms in the Lagrangian that give fermions their masses as well as their interactions is given by

$$\mathcal{L}_{yukawa} = h_u\overline{Q_L}\Phi u_R + h_d\Phi d_R + h_e\overline{\psi_L}\Phi e_R + h.c..$$

h_u, h_d and h_e are free parameters and are proportional to the relevant quarks and electron masses. Masses of all fermion constituents and vector bosons (the quanta that mediates forces) derive their mass values as a consequence Higgs phenomenon. This is the sense in which it is claimed that Higgs field, that pervades all space generates mass for the constituents in the universe.

We wish to point out that this is a bit of an exaggeration, given the fact the mass generated by the phenomenon yields very tiny values (2.15, 4.70 and 0.51 MeV) for u, d quarks and electron, which form almost all stable matter found in the universe. The bulk of mass for nucleons, in fact, arises from a different beautiful phenomenon and it is remarkable that this is a consequence again of the quantum principle. Recalling that spin (half integral angular momentum) of fermion has no classical analogue, we may assert that mass and spin are quantum attributes with little underpinnings in classical Physics.

3 Mass out of Nothing

We may suspend for a while the Higgs phenomena and deal with just strong and electromagnetic regime. The dynamics is governed by $SU(3)_C \times U(1)_{em}$ gauge theory that survives electroweak symmetry breaking. While electrodynamic forces govern the atomic structure of all elements and thereby all of chemistry, QCD is responsible to give us the nucleons and mesons as color neutral bound states of quarks with gluons as carriers of chromodynamic forces. Further, the residual (van der Walls like) interactions mimic the strong short range nuclear forces among nucleons and mesons build up the various nuclei, much like molecules are built out of electrically neutral atoms. The non-abelian gauge group is bestowed with asymptotic freedom (or vanishing coupling at very high frequencies or very short distances) and reciprocally confinement of color. Quarks and Gluons, that carry color quantum number are not to be seen as asymptotic states and are permanently confined within the color singlet modes. Mesons and Baryons as solutions of the dynamics constitute the spectrum of states. Their masses and the transition rates among them can be computed in QCD, just like QED provide *ab initio* atomic and molecular spectroscopy. Extreme precision in atomic spectroscopy and optics have been possible as a consequence of the development of high precision perturbative computation, since the small dimensionless coupling parameter $\alpha = 1/137$ renders reliability and order by order convergence of the computed quantities. In nuclear physics we do not have a small parameter to help us. However, in the underlying strong interaction, which we now recognize as emerging from QCD, it is possible to invoke perturbative QCD for a short distance (high Q^2) probe and use it to find both the scaling and quantitative scaling violation in deep inelastic scattering of leptons off protons and neutrons in the nuclei. This theory, however, is neither useful at predicting the low energy spectrum of baryons and mesons, nor determine the wave-functions of quarks in the hadrons. We need turn to non perturbative attempts to understand these features of QCD.

Lattice gauge theory reconstructs the theory on a space-time made up of lattice, such that as the lattice spacing is reduced and vanishes, the continuum theory is recovered. Methods of statistical mechanics are used to compute the various correlation functions and extract values for physical observables, such as masses of the bound states and resonant states, transition rates etc. of the theory, given just a few parameters that define the theory. We refer the reader to several review articles available (see [8]) and give here just an overview of what goes into the theory and the outcome thereof.

First, if there is no Higgs mechanism and no mass scales in the theory, how do we generate mass for the observed state? While classically the theory is scale invariant and hence has no mass parameter in the theory, when we deal the problem in quantum regime, a scale gets introduced as a process of regularization and renormalization. The coupling constant becomes a scale dependent parameter (Sidney Coleman called it a dimensional permutation [9]). Scale invariance now implies that as scale is changed there is a definite way all measured observables change. It also serves to shield intrinsic divergences, if any, in the theory to remain hidden in unobservable parameters of the theory. This provides us with a prescription to compute all measureable quantities in terms of a few parameters of the theory. QCD is defined with an intrinsic reference scale Λ_{QCD} at around 100 MeV, which we determine experimentally from the perturbative analysis of the deep inelastic scattering of leptons off nucleon targets [8].

The ingredients of theory is that we have a $SU(3)$ color gauge theory endowed with a fermion content made up of three generations of quarks. After the Higgs phenomenon we have quarks acquiring mass and phenomenological observation is that there are three light quarks and three heavy quarks. Of these u and d quarks are very light, c, b, t quarks are very heavy and s quark in the same order as Λ_{QCD}. We may begin with a toy model (Wilczek calls it QCD lite), setting all light quarks u, d and s massless and c, b and t infinite. The heavy flavours naturally decouple; and the three light massless quarks in the computation should give us a flavour $SU(3)$ spectra. Particle phenomenology of hadrons reflect an approximate flavour $SU(3)$ (with isospin I and hypercharge $Y(= B + S)$) symmetry, known to consist of a pseudoscalar meson octet ($\pi^{\pm}$, π^0, $K^{\pm}$, K^0, $\overline{K^0}$, and η), a vector meson nonet ($\rho^{\pm}$, ρ^0, $K^{*\pm}$, K^{*0}, $\overline{K^{*0}}$, ω and ϕ), a baryon octet (p, n, Λ, Σ^+, Σ^0, Σ^-,Ξ^0 and Ξ^-) and an excited baryon decimet (Δ^{++}, Δ^+, Δ^0, Δ^-, Σ^{*+}, Σ^{*0}, Σ^{*-}, Ξ^{*0}, Ξ^{*-} and Ω^-) as the prominent low energy spectra. Lattice gauge theory computations are able to quantitatively postdict this spectra. This computation has no input parameters, save the notion that $g_3(Q^2)$ depends on QCD scale Λ_{QCD}, which is obtained perturbatively from studying scaling violations of proton structure functions in deep inelastic scattering. In these experiments one uses weak and electromagnetic probes (e, μ, ν) to get the hadronic structure functions, whose Q^2 dependence gives us the scale of QCD. With only Λ_{QCD} as input, we make a statistical analysis of the system in a lattice framework of QCD. They yield a value $M_{p,n} \sim 890$ MeV, thus almost accounting for 95% of its mass as arising out of mass-less quark gluon radiation reaction.

4 Realistic hadron spectra from Lattice QCD

Lattice Gauge Theory aims to study Quantum Chromo Dynamics on a sufficiently large space-time (with periodic boundary conditions in all directions) regarded as a 4 dimensional grid of volume L^4 with lattice spacing of length a. A space time point x is specified by four integers through $x_\mu = n_\mu a$ and the limit $a \to 0$ and $L \to \infty$ lets us pass to continuum theory. Quark degrees of freedom $q_f(x)$, $f = u, d, ..$ reside on the lattice points and the gauge fields, gluons and photons, within $U_{\hat{\mu}}(x) = exp(i \int_x^{x+\hat{\mu}} G_{\hat{\mu}}(x')dx')$ on the links (numbering 8 for each site) that join a pair of neighbouring lattice sites. One then defines the partition function as the integral over all field variables of the Standard Model action of Gluons and Fermions;

$$S = S_G + S_F.$$

$$Z = \int DUD\psi D\overline{\psi}\ exp(-S[U, \psi, \overline{\psi}])$$

Statistical averaging of all possible configurations of the fields on the lattice allows us to simulate QCD and compute sampling of various field configurations on it. From these it is possible to extract experimentally measureable quantities. Powerful computational algorithms back up the effort to extract from it the outcome of the particle spectra and various transition amplitudes.

We saw in the preceding section computations with exact chiral invariance (with massless quarks) to obtain nucleon mass as 890 MeV. Next step is to let the parameters for quark masses $m_u = m_d$ and m_s free. In a Full QCD computation recently reported, BMW Collaboration, [10] used state of the art lattices with L/a = 64 and thus the space time has $N = 64^4 = 16,777,216$ sites. Computation involved matrices of dimension $12N \times 12N$ and storing about 4×10^{16} complex numbers. Adjoining figure illustrates the results. With pion (π), kaon (K) and cascade baryon (Ξ) masses as input values, we get the values of ρ, K^*, N, Λ, Σ, Δ, Σ^*, Ξ^* and Ω. We have a remarkable agreement in the description of the observed set of pseudo scalar mesons, vector mesons, baryons and excited baryon states. Nucleon is, as expected, at 940 MeV.

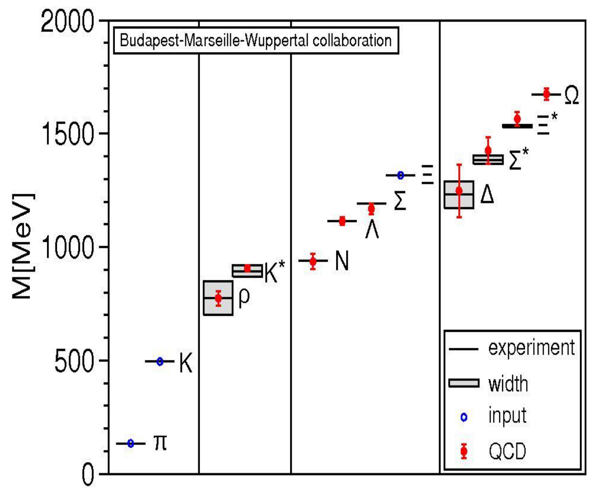


Figure 3: Spectrum of low energy spectra, computed from first principles in QCD; source [10]

Lattice characterization of QCD, should not be seen as an approximation to continuum space-time, but (generically) an unavoidable interim part in the definition of the theory. The procedure is intrinsically gauge invariant since it deals with gauge invariant content all the time, unlike in a perturbative treatment, where a choice of a gauge has to be made and care must be exercised to ensure that the final outcome is gauge invariant. This Lagarangian - regulated renormalized extrapolation - respects confinement of color, chiral symmetry and scale invariance limits adequately. We are able to address many features of QCD in the low energy regime of hadron physics that are available for experimental study; such as decay constants f_π, f_K; semileptonic form factors that are appropriate for computing $B \rightarrow D l\nu, K l\nu, \pi l\nu$; $l = e, \mu$ etc. To begin with one computes in the so called quenched approximation, in which quark degrees of freedom are ignored in order to keep the 'cost' of computation in terms of available computing resources kept within manageable level. As tera and penta flop speed in computing get developed more ambitious project of 'full' QCD are possible as a major global collaborative endeavour. We will paraphrase Wilczek[4] in identifying the conceptual roots that shaped the outcome as represented in the fig. 2. Special theory of Relativity appears to demand that the interactions are local; the local interaction bring in fields with arbitrarily large frequencies (energy) and short wavelengths (high momenta) that may cause divergence that will render the calculations unreliable. Non abelian gauge theories avoid it, by virtue of the property of asymptotic freedom that weakens the coupling of the dangerous modes. This happy result occurs only for the gauge invariant minimal couplings as are considered in these exercises.

QCD, a gauge theory based on gauge group $SU(3)$ color triplet quarks and color octet gluons, both degrees of freedom remaining confined in gauge singlet hadrons is highly constrained, supporting very few free parameters. A mass parameter for each flavor quark (together with flavor mixing angles of Cabibbo -Kobayashi Masakawa matrix) and just an overall coupling strength is all that one is allowed. Since asymptotic single quark states are never seen, the mass parameters of quarks are to be seen as just inputs that figure in getting the masses of hadrons. The coupling $\alpha_s(Q^2) = g_3(Q^2)/4\pi\hbar c$ is large when Q is less than or of order Λ_{QCD} and fluctuations in gluon dominates the dynamics. Bulk of the nucleon mass, we may presume, thus gets built on a tiny chiral symmetry breaking mass of u and d quarks by the gluon dressings carrying most of energy associated with the state. This is reminiscent of what was indeed an old speculation of Lorentz as the origin of electron mass. He associated rest-mass/energy of the electron with the energy in the form of Electric field residing in the space, $1/(2\epsilon_0)\int d^3x E^2(x)$, sort of radiation reaction on the motion of the electron. For a point charge this will be indeed divergent, but is finite for an electron with a distributed size of range $\lambda = \hbar/mc$, its Compton wavelength. We may use this to fix the radius of electron (which turns out be of order $\alpha\lambda$). While this is not anymore regarded as the origin of electron mass, (now that α is no longer a fixed constant and the Higgs coupling rather than finite size of electron as dictating it) we find that the hadron masses seem to possess some features of gluonic radiation reaction as generating bulk of the mass, in a somewhat similar picture as that of Lorentz. It is remarkable that Lattice framework of QCD provides dependable *ab initio* prediction, thanks to high speed computing resources and very smart dedicated algorithms available now for such a computation.

5 Summary

While classically mass is an extrinsic kinematic parameter that signifies the amount of matter, quantum regime makes *mass* a dynamical input. This feature for mass has two somewhat different origins. First, we observe that Quantum Chromo Dynamics (QCD), that governs interaction among quarks and gluons, is responsible for the mass to primary nucleons, the most abundant source of visible matter in the universe. Next, apart from the u and d quarks (and the leptons ν_e and e) we need at least two more generation of quarks and leptons to complete the matter content. All of them were abundant and in a dynamical equilibrium at the very early stages of the universe, but now most (except quarks and leptons of the first generation) are only seen as short lived intermediate particles. These as well as the weak interaction-mediating bosons $W^{\pm}$ and Z get their mass as a result of the coupling with the Higgs scalar field. We emphasize that both features point to the notion that *mass* is essentially a quantum consequence, which has no classical analogue.

References

[1] G. Aad et al. (The ATLAS Collaboration), *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*," arXiv:1207.7214 (hep-ex)

[2] S. Chatrchyan et al. (The CMS Collaboration), *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC* arXiv:1207.7235 (hep-ex)

[3] **God particle created by Big Bang 13.7 bn years ago seen**. The presentation shown at PU Auditorium, webcast from CERN shed light on the existence of Higgs particle, which gives *mass to all* the existing matter in the visible universe. Times of India July 5 2012

[4] F Wilczek. *The origin of mass*, Invited Review for Central European Journal of Physics arXiv:1206.7114 [hep-ph] 2012

[5] Y Aharanov and D Bohm. Phys. Rev. 115 485 (1959)

[6] D Gross and F Wilczek. Phys. Rev. Lett. 30 1343 (1973)

[7] H D Politzer. Phys. Rev. Lett. 30 1346 (1973)

[8] J Beringer *et al* PDG Phys Rev. R D86, 010001 (2012)(http://pdg.lbl.gov)

[9] S Coleman and E Weinberg. Phys. Rev. D7 1888 (1973)

[10] S Durr, Z Fodor, J Frison, C Hoebling, R Hoffman, S D Katz, S Krieg, T Kurth, L Lelbouch, T Lippert, K K Szabo and G Vulvert. Science 322 (2008)

Estimating the Pressure of a Gas in a Balloon and Motion through Resistive Media

Sanjay Harrison[1] and Sindhu Vincent [2]
Northern Deanery. UK
47 Annand Way, Newton Aycliffe, Co Durham, DL5 4ZD, United Kingdom
quibit21@hotmail.com[1], sindhuvincent@gmail.com[2]

(*Submitted* June 13 2012)

Abstract

A simple method to estimate the pressure of a gas (P) inside a balloon is described. The basis of this method is derived from the principle of floatation – a principle most high school students would be familiar with. In this article we also describe the dynamics of the gas filled balloon as it traverses the depth of the liquid in the vessel. The dynamics of the balloon is described first by neglecting the effects of resistance through the liquid. The effect of resistance or drag is taken into account in subsequent sections of this paper. Expressions for the pressure of the gas in the balloon, the velocity (v) and acceleration (a) at a given instant and the terminal velocity ($v_{terminal}$) of the balloon are derived.

Introduction

Simple phenomena can very often hide some very interesting physics and studying such phenomena can lead to a greater appreciation of how seemingly disparate branches of physics come together and give rise to the phenomena in question. This paper elucidates the physics of linear motion through a resistive media and demonstrates how it can be used to determine estimates of gas pressures. The simple experiment described brings together kinematics and thermodynamics. It is believed that the demonstration of such connections would lead to greater appreciation amongst younger readers in particular of general physical principles.

Case 1 : Resistance due to Medium Ignored

The balloon containing the gas whose pressure is to be determined is immersed in a liquid (eg water) whose density is known (d_w) and held there with the help of an angled thin rod. The depth to which it is immersed (h) is measured and the moment it is released, a stopwatch is started and the time (t) the balloon takes to reach the surface when released is measured.

In calculating the pressure of the gas within the balloon, the following simplifying assumptions are made;

- The resistance experienced by the balloon as it makes its way to the surface of the liquid is negligible. This is justified by noting that the resistive force is proportional to the velocity of the balloon at any given instant. As the balloon only travels the depth of the vessel, the velocity acquired would be small.
- The weight of the balloon is negligible compared to the up thrust (U). This can be easily seen by noting that the numerical value of the volume of the balloon is much greater than the mass of the balloon.
- The gas within the balloon behaves as an ideal gas. This assumption is primarily for the purpose of calculational ease. It has to be emphasised that most gases only exhibit ideal behaviour at high temperatures and low pressures. However, as is evident

from the title of this paper, the experimental method described provides an 'estimate' of the pressure of the gas within the balloon. While this experimental method may not provide an exact value for the pressure, we believe that it in no way lessens the educational value it provides to the reader by connecting seemingly disparate areas of physics.

The mass of the gas (M) in the balloon will experience an upward force when immersed in the water containing vessel giving it an upward acceleration (a). Acceleration due to gravity is represented by 'g'. Once the relevant measurements have been made, the density of the gas (d_g) can be estimated as follows;

$$M\,a = V\,d_w\,g - M\,g$$

$$V\,d_g\,a = V\,d_w\,g - V\,d_g\,g$$

$$a = \frac{d_w}{d_g} - 1 \quad g$$

Applying the second equation of motion to the balloon travelling the depth of the vessel (h);

$$h = \frac{at^2}{2}$$

$$h = \frac{gt^2}{2} \quad \frac{d_w}{d_g} - 1$$

$$d_g = \frac{gt^2}{2}\,\frac{d_w}{2h+gt^2} \qquad (1)$$

Once the density of the gas has been calculated from the above equation, the ideal gas equation is modified as follows (P = pressure, V = volume, n = number of moles, M = mass of gas, m = molar mass of gas, T = temperature of gas, R = gas constant);

$$P\,V = n\,R\,T$$

$$\frac{PM}{d_g} = \frac{MRT}{m}$$

$$P = \frac{d_g}{m}RT \qquad (2)$$

Substituting equation 1 in equation 2 we obtain;

$$P = \frac{gt^2}{2} \cdot \frac{d_w}{2h+gt^2} \cdot \frac{RT}{m}$$

The temperature of the gas is assumed to be the same as the room temperature. The method described above can be used even if the gas filled balloon is suspended in a denser gas (eg helium filled balloon in air).

Case 2 : Resistance of liquid taken into account

This part of the theory has an extra level of complexity to it and it will be shown that an expression for the acceleration is only obtained in a transcendental form. Therefore a modification of the experimental method is required. We start by looking at the forces acting on the balloon as it travels through the liquid. We obtain,

$$M\,a = U - Mg - \frac{v^2 d_w AC}{2} \qquad (3)$$

The last term in equation 3 is the drag or resistance force experienced by the balloon as a result of its motion through the liquid. This expression for the drag is known as the drag equation[1] where 'A' is the cross sectional area of the balloon (which is taken as circular) and 'C' is the drag coefficient (0.47). It must be noted that the drag is proportional to the square of the velocity and therefore when the balloon is stationary the drag is equal to zero. This fact is used to modify the experimental setup.

The balloon is immersed fully in the liquid, but this time a sensitive spring balance is inverted and attached to the base of the balloon. The reading obtained would be equal to the upward force exerted on the balloon when it is stationary or just before it is released. Equation 3 then reduces to;

$$M\,a = U - M\,g \qquad (4)$$

Substituting the volume and respective densities into the above equation we get;

$$V\,d_g\,a = V\,d_w\,g - V\,d_g\,g$$

This simplifies to;

$$a = \frac{d_w}{d_g} - 1 \quad g$$

As this is the acceleration of the balloon, the force acting on the balloon which can read off the spring balance is;

$$F = Vd_g \quad \frac{d_w}{d_g} - 1 \quad g$$

From this equation, an expression for d_g can be obtained as;

$$d_g = d_w - \frac{F}{Vg}$$

This expression for the density of the gas can be substituted into equation 2 to obtain an expression for the pressure of the gas in the balloon. ie

$$P = d_w - \frac{F}{Vg} \quad \frac{RT}{m} \qquad (5)$$

Dynamics of the Balloon with Resistive Forces:

The equation of motion of the balloon is given by equation 3;

$$M\,a = U - Mg - \frac{v^2 d_w AC}{2}$$

Replacing the mass terms with the product of volume and the respective densities and expressing the volume of the balloon in terms of its cross sectional area (A), subsequent rearrangement results in;

$$a = \frac{d_w}{d_g} - 1 \quad g - \frac{3v^2 C d_w}{8rd_g}$$

where 'r' is the radius of the spherical balloon. As acceleration is defined as the rate of change of velocity, the above equation can be simplified to give;

$$\frac{dv}{dt} = \frac{d_w}{d_g} - 1 \quad g - \frac{3Cd_w}{8rd_g} \quad v^2 \qquad (6)$$

This differential equation is of the form;

$$\frac{dv}{dt} = A - Bv^2 \qquad (7)$$

Where $A = \frac{d_w}{d_g} - 1 \quad g$ and $B = \frac{3Cd_w}{8rd_g}$.

If we define $p^2 = \frac{A}{B}$ and given that v = 0 at t = 0, equation 7 can be solved by the method of partial fractions to give;

$$v = \frac{1 - e^{-2Bpt}}{1 + e^{-2Bpt}} \quad p \qquad (8)$$

This expression gives the velocity of the balloon at any instant of time. Differentiating this equation with respect to time we obtain the acceleration;

$$a = \frac{2Bp^2 e^{-2Bpt}}{1 + e^{-2Bpt}} - \frac{2Bp^2 e^{-2Bpt} \quad 1 - e^{-2Bpt}}{1 + e^{-2Bpt} \ ^2} \qquad (9)$$

The terminal velocity is the constant velocity acquired by the balloon when the resistive forces equal the upward acting force on the balloon. The terminal velocity can be obtained from equation 8 by taking the limit of the expression as 't' tends to infinity. This would give;

$$v_{terminal} = p$$

From the definition of p, we obtain;

$$v_{terminal} = \sqrt{\frac{8rg\,(d_w - d_g)}{3Cd_w}}$$

Taking the limit of equation 9 as 't' tends to infinity we note that

$$a = 0$$

This is accordance with the balloon acquiring a 'constant' terminal velocity.

Conclusion:

This paper describes a simple experimental setup that can be easily replicated in the classroom for determining the pressure of a gas inside a balloon. The experiment is described and adapted for when resistive forces are taken into account. In addition, we work out the dynamics of the balloon as it moves through the resistive medium and obtain expressions for its velocity, acceleration and terminal velocity. Since the theory of the method described ties together seemingly disparate areas of physics such as floatation, kinematics and gas laws, it can be used to illustrate to students, the wide applicability of physics and how it can be used to solve problems with very basic equipment. This could potentially stimulate interest among students and promote learning.

Reference

1. Batchelor, G.K. (1967). *An Introduction to Fluid Dynamics*. Cambridge University Press. ISBN 0521663962.