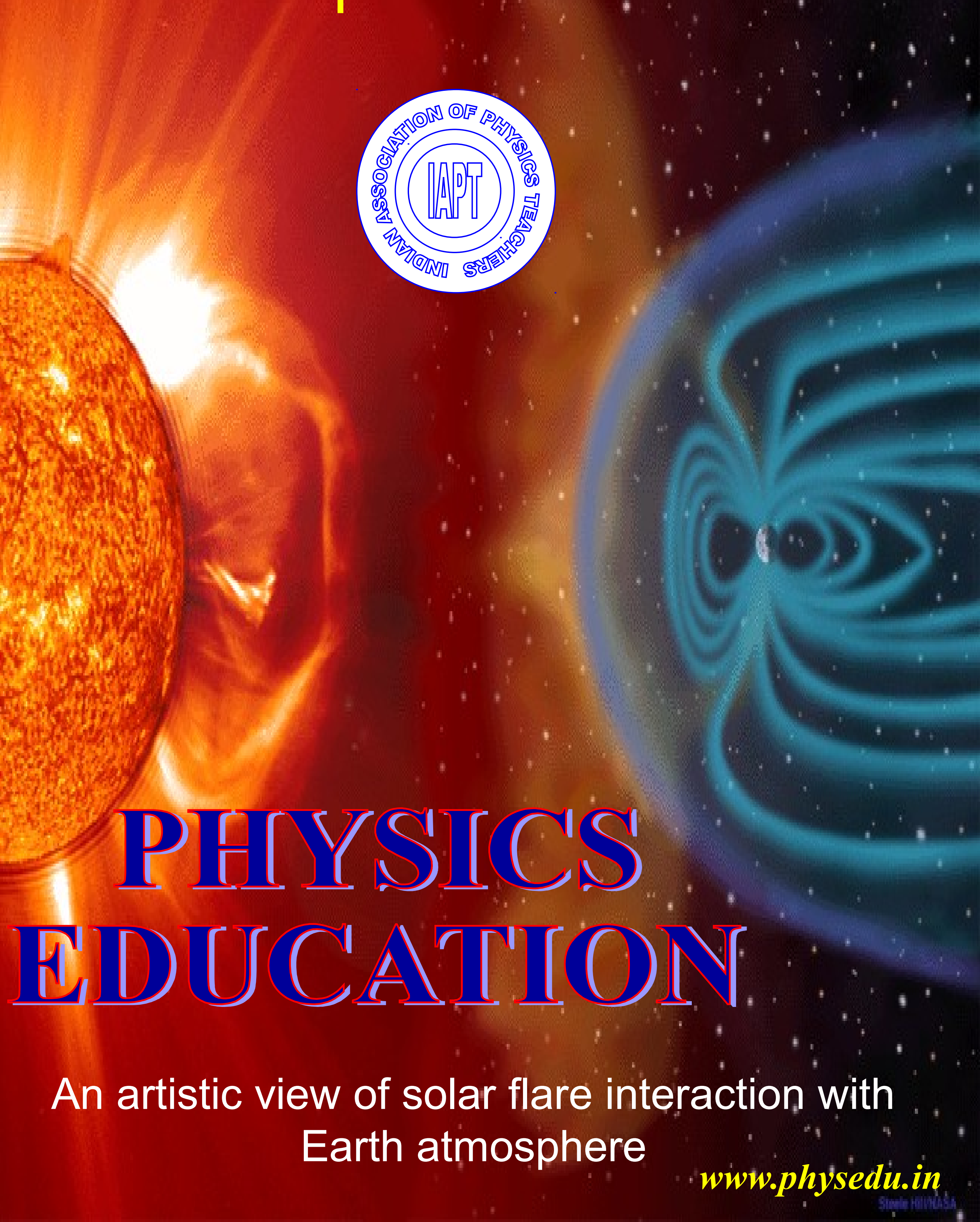
Vol 31 No. 2 Apr - Jun 2015 ISSN 0970-5953



INDIAN ASSOCIATION OF PHYSICS TEACHERS

IAPT

PHYSICS EDUCATION

An artistic view of solar flare interaction with Earth atmosphere

www.physedu.in

Volume 31, Number 2

In this Issue

EDITORIAL

It is a pleasure to publish the issue 31.2 of Physics Education containing a wide range of interesting articles. C K Raju's article is of fundamental nature, giving a formulation of theory of gravitation in terms FDEs. P. Mandal gives an interesting account of trapping of charged particles, while Sazonov gives a novel technique to calculate the contribution of magnetic energy to the energy of the field of a rotating metal sphere. H. S. Rawal deals with a problem of quark confinement in simple terms. The article on Foucault pendulum by S. S. Verma includes the problems of constructing one. The article on Solar flares by Kumar, Bhatt and Jain gives insights into a new field for students. Minkin and Shapovalov show subtle implications of physical situations for vector addition. A pedagogical formulation of the requirement of complex numbers in quantum mechanics is given by Maynard, Lambert and Deering.

I wish you a very happy reading!

Pramod S. Joag.

Chief Editor, Physics Education
Chief-editor@physedu.in,
pramod@physics.unipune.ac.in

Functional differential equations. 4: Retarded gravitation

C. K. Raju

ckr@ckraju.net

(Submitted 27-06-2015)

Abstract

Are functional differential equations (FDEs) only about electrodynamics? No. They apply also to gravitation. We explain a recent reformulation of gravitation, called retarded gravitation theory (RGT), which is Lorentz covariant, and uses functional differential equations. RGT modifies the Newtonian "inverse square law" gravitational force: the RGT force depends upon (a) retarded distance, and (b) includes a velocity-dependent term. RGT, since Lorentz covariant, theoretically improves on Newtonian gravitation. At the same time, RGT has the practical advantage over general relativity theory (GRT) that a solution of the many-body problem is feasible in RGT. Hence, RGT can and ought to be applied to the galaxy where Newtonian physics apparently fails but GRT cannot be applied. The tiny velocity dependence of the RGT force is amplified across a hundred billion co-rotating stars in the galaxy, so that non-Newtonian velocities of stars in spiral galaxies are to be expected on RGT, even without dark matter. Possible experimental tests of RGT include the flyby anomaly observed for NASA spacecraft which depends systematically on velocity-effects due to the rotation of the earth.
We further clarify that Laplace's objection to pre-relativistic naive theories of retarded gravitation (NRG) does not apply to RGT. We solve the 2-body FDEs of RGT for the sun-Jupiter case: the system is stable despite tiny differences from Newtonian gravitation. Thus, FDEs are a general feature of post-relativity physics.

1 Recap

In three earlier articles[1, 2, 3] in this series, we saw that functional differential equations (FDEs) are fundamentally different from ordinary differential equations (ODEs). Hence, doing physics with FDEs leads to a paradigm shift in physics. Further, FDEs arise naturally in classical electrodynamics: *without* any new physical hypotheses but just by doing the math right. The right way to solve for the classical hydrogen atom, even without radiation damping, is to use FDEs and that changes the qualitative features of the solution.

What happens if we have radiation damping? The problem of the motion of even a single charged particle, in classical electrodynamics, with radiation damping has remained mathematically unsolved for a century because of runaways. These runaways can be controlled by modifying Maxwell's equations at the microphysical level, so that the equations of motion of even a single charged particle become FDEs.

Before proceeding further to quantum mechanics, there is one doubt which needs to be settled. Are FDEs only about classical electrodynamics? No. They are about resolving a fundamental conceptual flaw in Newtonian physics. I have dealt with this issue of Newtonian physics in detail in previous articles in this very journal,[4, chp. 2, chp. 3a, chp. 3b] and will only summarise the key points here.

2 The problem of time in Newtonian physics

Consider Newton's first law of motion. It states that, in the absence of external forces, a body continues in its state of rest or uniform motion. Is this meaningful? It is easy to understand "rest", but what is "uniform motion"? A body is said to be in uniform motion if it covers equal distances in equal times. But what are equal times?

When we say that one hour in the past is equal to one hour in the future, there is no way to verify it empirically. Obviously, we cannot bring back one hour in the past and compare it in the present with one hour in the future. We must use a clock. But, clocks differ, so *which* clock should one use? Uniform motion according to my heart beats would not be uniform motion according to a simple pendulum, and vice versa. Without a definition of equal intervals of time, or the definition of an "ideal clock", there is no basis on which to say that a mechanical clock is "better" than heart beats.

So, what exactly is an ideal clock? In his *Principia*, Newton admitted that days and nights are unequal, as are the swings of a pendulum, and that no natural phenomenon would provide an ideal clock. But he reached the strange conclusion that it was unnecessary to define equal intervals of time. He said that he was concerned only with "absolute, true, and mathematical time, which flows on without regard to anything external". Each of these adjectives: "**absolute**", "**true**", "**mathematical**", and "**without regard to anything external**" makes clear that Newton took time as an aspect of metaphysics. In short, he thought it was all right if God knew what equal intervals of time were, even if humans did not.

This was a mistake because to do physics, humans too need to know what equal intervals of time are. Indeed, Newton's predecessor and mentor Barrow had emphasized the need for a clear physical definition of equal intervals of time, saying those who did physics without it were "quacks".[5] *Why* did Newton make time metaphysical? Newton thought that making time metaphysical was the way to make "perfect" the notion of $\frac{d}{dt}$ needed for his second law. This related to the European misunderstanding of the Indian calculus imported into Europe in the 16th c. This is an interesting but long story, which I have told elsewhere.[6, 7]

For common applications of Newtonian mechanics, to planetary motion and ballistics, many common clocks "work". However, Newton's failure to provide a physical definition of equal intervals of time, became prominent during attempts to reconcile electrodynamics with Newtonian physics at the turn of the 19th c. The solution provided by relativity was to define equal times in a way which preserved electrodynamics but required a modification of Newtonian physics.

Physics texts teach relativity differently: they teach that relativity began with the Michelson-Morley experiment which proved the absence of ether and the constancy of the speed of light. That, however, is not correct: one cannot measure the speed of light or anything else without a clock, and a positive result in the experiment (as later found by Miller) is no evidence either for ether or for a varying speed of light.[4, p. 56–57] In fact, as explained in an earlier article in this journal, the Michelson-Morley experiment was NOT designed to test the existence of ether: it was designed to test between the two ether theories of Fresnel and Stokes. Amusingly, it came out in favour of the wrong theory: the Stokes theory, which involved a mathematical absurdity. Hence, Lorentz thought it was preferable to believe that the arm of the Michelson interferometer contracted in the direction of motion.

Now, a clock is required even to measure *lengths*: for a moving rod, one must note the positions of both ends of the rod *simultaneously*, and simultaneity is decided by a clock. If one postulates that the speed of light is constant, then a photon bouncing between two parallel mirrors marks equal times between bounces, and this provides an ideal clock. The Lorentz-Fitzgerald length contraction is a natural consequence of using such a clock.

Note clearly that the constancy of the speed of light is a postulate, *not* an experimental fact. This postulate of constant speed of light automatically leads to the Lorentz transform which Poincaré derived and so named. That is, the special theory of relativity came about as the solution to the problem of equal intervals of time in Newtonian physics.

As for "ether", the word is confusingly used in multiple senses. One sense is as an absolute reference frame. But the original sense of ether (= sky = *ākāṣa*, as in the Vaiśeṣika sūtra) relates to action by contact (*saṃyoga*). Eliminating ether also eliminates action by contact, and admits, for example, delayed action at a distance. This is mathematically equivalent to replacing ODEs by FDEs (which Poincaré called "equations of finite differences" [4, p. 116]). Einstein, to whom special relativity is usually attributed, never understood this point, for he mistakenly kept approximating FDEs by ODEs, until late in his life.[4, p. 122]

This *process* of development of relativity, by identifying and resolving a conceptual flaw in Newtonian physics, as well as the connection of relativity with FDEs, are both obscured by usual accounts of the theory of relativity which focus on glorifying an individual, Einstein. (It is on record that Einstein knew of Poincaré's work until 1902. In his 1905 special relativity paper, he casually used the strange term "longitudinal mass" first circumspectly used by Lorentz in 1904. Einstein also used the novel term "relativity" first used by Poincaré in his 1904 paper (instead of his earlier "principle of relative motion"). Einstein later denied reading both the 1904 papers, of Lorentz and Poincareé, and his 1905 paper on (special) relativity cites absolutely no references.)

3 Modifying gravitation

Special relativity modified Newton's laws of motion; but that is not enough, Newtonian gravitation too must be modified for the two come as a package deal. Newtonian gravitation involves instantaneous action at a distance which is incompatible with special relativity, where the speed of light is a limiting speed. The general theory of relativity (GRT) did modify Newtonian gravitation. However, GRT is enormously complicated: in a century since GRT was formulated, even the two body problem could not be solved in it. This creates a peculiar problem as follows.

3.1 Galactic rotation curves

Newtonian gravitation worked well for the solar system, but it fails for the galaxy. In the solar system the rotational speed of a planet of mass M_p is determined by

$$\frac{GM_S M_p}{r^2} = \frac{M_p v^2}{r} \tag{1}$$

where M_S is the mass of the sun, v is the rotational velocity of the planet round the sun, and r is the distance of the planet from the sun. This means that the rotational velocities of planets $v \propto \frac{1}{\sqrt{r}}$ decline with distance r from the centre. Or, in terms of the time period $T = \frac{2\pi r}{v}$, we must have $T \propto r\sqrt{r}$. This accurately fits observations: Pluto some 39.5 times more distant from the sun than earth takes $39.5 \times \sqrt{39.5} \approx 248$ earth years to complete an orbit round the sun.

However, what happens in a spiral galaxy is starkly different. In spiral galaxies, the rotational velocities of stars, instead of declining, are observed to *increase* as one moves out from the centre. (Fig. 1) This is contrary to what one expects from Newtonian gravitation.

3.2 Dark matter

Of course, the Newtonian theory can be easily "saved" by supposing that there is invisible dark matter (DM) in the galaxy. Perhaps that is so: but at least we expect a decline in rotational velocities of

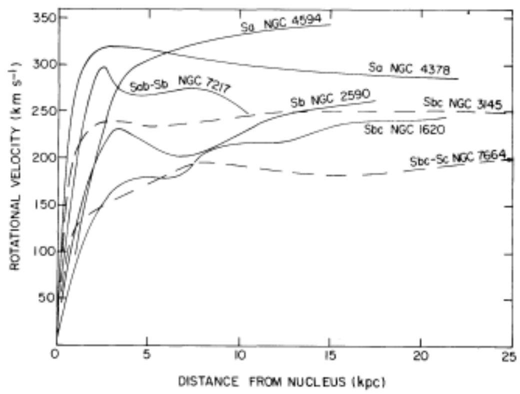


Figure 1: Rotation curves of various galaxies. The rotational velocities increase and then become roughly constant.[8] This is contrary to the expected behaviour on Newtonian gravitation that, sufficiently far from the nucleus, rotational velocities must decline as $\frac{1}{\sqrt{r}}$ with distance r from the nucleus.

stars as we move to the edge of the visible galaxy. Unfortunately, even that expectation is belied. It is clear from Fig. 1 that the rotational velocities of stars, instead of declining, become approximately *constant* at the edge of the galaxy. Therefore, to "save" the theory we must make one more hypothesis: that the hypothetical invisible dark matter is distributed in a peculiar way in the form of a halo round the galaxy, with its density reaching a peak where the luminous matter thins out to zero.

Now why should that be so? The hypothetical, invisible dark matter, whatever its composition, has exactly the same gravitational properties as the luminous matter in galaxies. On the scale of the galaxy, gravitation is the dominant force which decides structure. So why should luminous matter and dark matter be distributed in such strikingly different ways? No clear explanation has emerged so far.

3.3 MOND

Dissatisfaction with the DM hypothesis led to the formulation of another theory: modified Newtonian dynamics (MOND).[9, 10] In its original form, the theory simply supposed that the gravitational force law itself changed at the scale of the galaxy on the phenomenological grounds of observations. Moreover, it did not initially correct what we now know to be a critical conceptual defect in Newtonian physics.

Could GRT explain any part of the departure from Newtonian gravitation? To answer this we need to be able to apply GRT to the galaxy. Unfortunately, that is not feasible: GRT is too complex to be used to solve the many body problem of a galaxy typically involving hundreds of billions of stars. Even modelling a collection of discrete objects is very difficult in GRT. Therefore, the only thing available is to fall back on Newtonian gravitation just believing it to be a good approximation to GRT at those scales.

4 RGT

This situation motivates retarded gravitation theory.[7] We know that special relativity is an essential conceptual correction to Newtonian physics. Can we have a theory of gravitation compatible with special relativity? Poincaré did attempt to find such a theory (with a different motivation) but did not fix on a definite expression for the force, or try to solve the problem of galactic rotation curves, which was not known in his time.

Some people might ask: why look for such a theory when we already have the "ultimate" theory, namely GRT? One simple answer is this: it is no use having an ultimate theory which is not usable! It is like saying ultimately "God knows everything", but we have no way to read the mind of God! In the context of the galaxy, the theory which is actually used is Newtonian gravitation. RGT, being a Lorentz covariant theory of gravitation, improves on that. As we will see, RGT does help us to bypass the additional hypotheses introduced by both DM and MOND.

4.1 Derivation of the expression for the force

How does one make gravitation Lorentz covariant? The derivation of the Lorentz-covariant gravitational force is so simple that it can be reproduced in its entirety here.

We start with a reference frame in which the test particle ("attracted body") is a mass point (at rest) at the origin. The "attracting body" is located at the retarded position described by the 4-vector $X = (ct, \vec{x})$ and moving with a 4-velocity $V = \gamma_v(c, \vec{v})$, both at *retarded* time $t = -\frac{r}{c}$. Here, $\vec{x} = (x, y, z)$, $r = \sqrt{x^2 + y^2 + z^2}$, and $\gamma_v = (1 - \frac{v^2}{c^2})^{-\frac{1}{2}}$ is the Lorentz factor.

Let $F = (T, \vec{f})$ be the 4-force experienced by the attracted body. This 4-vector transforms in the same way as the 4-vectors X and V, so we take it to be given by a linear combination

$$F = aX + bV, \tag{2}$$

where a, and b are Lorentz invariants to be determined. Since a and b are Lorentz *invariant*, the expression (2) for the 4-force F would be Lorentz covariant, as required.

For the case where the attracting body is also at rest ($\vec{v} = 0$), we require that the 3-force must approximately agree with the Newtonian gravitational force $\vec{f} = k(\frac{x}{r^3}, \frac{y}{r^3}, \frac{z}{r^3})$, where $k = Gm_0m_1$, the two (rest) masses are m_0 and m_1, and G is the Newtonian gravitational constant. (Note that the sign conventions we are using are the opposite of the usual ones, since the "attracting body" is at X, and the force is in the direction of its retarded position.) Therefore, $a \approx \frac{k}{r^3}$. This suggests that $a = -\frac{kc^3}{a_1^3}$ where a_1 is the Lorentz invariant quantity $a_1 = X.V = \gamma_v(c^2t - \vec{x}.\vec{v})$, which equals $-cr$ when $\vec{v} = 0$, and approximately equals $-cr$ when $v = ||\vec{v}||$ is small compared to c. That is,

$$a = -\frac{kc^3}{(X.V)^3} \approx \frac{k}{r^3}. \tag{3}$$

We now use the fact that the components of the 4-force are not independent, but must satisfy [11]

$$F.U = 0, \tag{4}$$

where $U = \gamma_u(c, \vec{u})$ is the 4-velocity of the particle on which the force acts. This comes about simply since the revised form of the equations of motion is now

$$m_0 \frac{d^2Y}{ds^2} = F, \tag{5}$$

where m_0 is the rest mass and s is proper time along the world line, $Y(s)$, of the "attracted particle". Since the 4-force F is parallel to the 4-acceleration of the particle on which it acts, it must be perpendicular to its 4-velocity U (which is a vector of constant norm). Accordingly, taking the dot product of U with both sides of (2), we obtain

$$0 = a(X.U) + b(V.U). \tag{6}$$

Now the dot products $X.U$ and $V.U$ are scalars, or Lorentz invariants, and the Lorentz invariant a is already determined. Hence, (6) determines b as a Lorentz invariant. Explicitly,

$$b = -\frac{a(X.U)}{(V.U)} \approx \frac{k}{cr^2}. \tag{7}$$

Note that we would not have been able to satisfy the requirement (4) had we already set $b = 0$ to begin with. This shows that the Lorentz covariant gravitational force we seek *cannot* be purely position dependent but must depend also on velocity.

Substituting these values of a and b in (2), the force in RGT is explicitly given by

$$F = -\frac{kc^3}{(X.V)^3}X + \frac{kc^3}{(X.V)^3}\frac{(X.U)}{(V.U)}V. \tag{8}$$

Since the equations of motion (5), and the expression for the force (8) are Lorentz covariant, we can use these expressions in any Galilean frame, and are not tied to any special frame. Note, however, that RGT, unlike GRT, is restricted to Galilean frames.

In studying motions such as those of stars in the galaxy we can use the non-relativistic approximate expressions for a and b given in (3) and (7). This leads to

$$F \approx \frac{k}{r^2}\left(\frac{X}{r} + \frac{V}{c},\right), \tag{9}$$

which simple form exhibits clearly the departure from Newtonian gravitation.

Thus, we have made two changes to Newton's "inverse square law" of gravitation. First, the RGT gravitational force uses the retarded distance, not the instantaneous distance between the two bodies. Second, the gravitational force cannot be a pure "inverse square law" force (even with the retarded distance), but has a velocity-dependent ($\frac{v}{c}$) component. This RGT modification of Newtonian gravitation is completely different from other earlier modifications of the "inverse square law". Furthermore, it is not *ad hoc* or speculative like earlier modifications; rather RGT is a logical consequence of an *essential* correction to a conceptual defect about time in Newtonian physics.

5 The solution for the galaxy

Now special relativistic effects, such as the $\frac{v}{c}$ term in the above force, are believed to be relevant only when velocities approach that of light. However, this piece of text-book wisdom is true only for the *one* body problem, the only problem solved by texts in special relativity. For the galaxy, however, we need to do a many-body problem.

Now, stars in a spiral galaxy all *systematically* co-rotate in one direction. Could a tiny but systematic $\frac{v}{c}$ effect become significant when summed over a large number of stars? Specifically, The observed rotation velocities of stars in spiral galaxies are of the order of a few hundred km s^{-1} corresponding to $\frac{v}{c} \sim 10^{-3}$, which is small. However, a systematic effect of this order must be summed over some 10^{11} stars in a galaxy. Could the sum be significant?

To make a quick check we can re-frame the question. Suppose we have numerous mass points spread in a disk rotating around a central mass. Suppose, now, we introduce a test mass into this configuration, and let us further suppose the test mass is moving with the Newtonian velocity required for equilibrium at a distance r from the centre. What will happen to the test mass?

For a test mass, the calculation is very simple, for the meaning of a "test" mass is that we neglect the effect of that mass on the remaining particles (the galaxy). That is we prescribe the motion of the other particles not only in the past, but for all time. Under these circumstances of a one-body problem, where the motion of all remaining bodies is prescribed for all time, the FDEs of motion in RGT reduce to ODEs, which can be readily solved. However, the force still differs from that of Newtonian gravitation.

According to Newtonian gravitation the rotational *velocities* of the other masses are irrelevant, and only the total mass counts. Hence, our test mass which begins in Newtonian equilibrium, should continue in equilibrium. On RGT, however, the test particle is violently accelerated. Depending upon the total mass it may stay within the system, with a non-Newtonian velocity, or get thrown out. Therefore, on RGT, a large number of co-rotating particles can significantly increase the rotational velocity of a test particle in Newtonian equilibrium. Further, unlike Newtonian gravitation, if we consider a shell of rotating particles, the velocity effect acts on the test particle even inside the shell.(Fig. 2)

The story has a very important moral: tiny special relativistic effects, at non-relativistic velocities, can add up across a large number of particles, and become immense. The text book claim that special relativity matters only at relativistic velocities needs to be corrected: that claim is true only for the one body problem.

Thus, RGT predicts that stars in a spiral galaxy will have non-Newtonian velocities just because the velocity-dependent gravitational force adds up across a large number of co-rotating stars. It is not necessary to hypothesize dark matter just to explain non-Newtonian velocities in spiral galaxies, but if there is any dark matter, its effects would be in addition to those predicted by RGT.

What about the other feature of rotation curves that rotational velocities become constant at the edge of the galaxy? In principle, this feature too admits a simple explanation in RGT. Far from the centre of the galaxy, the gravitational pull of the central mass becomes weak, and the velocity effect becomes more prominent. Consider two nearby stars co-rotating at the edge of the galaxy. The velocity dependent component of the RGT gravitational force will tend to equalise their velocities. Thus, there is a simple and natural explanation for the approximate constancy of

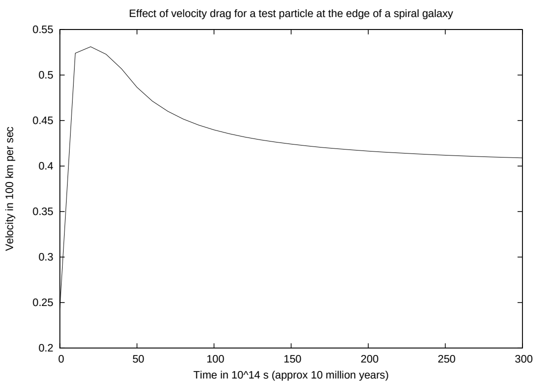


Figure 2: **The velocity effect of retarded gravitation:** The velocity of a test particle in our model galaxy increases due to velocity drag, and the particle escapes. The plot is velocity vs time. Time units are approximately 10 million years, while length units are 1 kpc. A central mass of 1.5×10^{10} solar masses is surrounded by 10,000 particles (each of 10^5 solar masses) in a rotating ring of radius 12 kpc. The test particle is initially in Newtonian equilibrium at 12.2 kpc.

star velocities at the edge of the galaxy, and there is no need to hypothesize halos of dark matter.

Thus, the hitherto mysterious qualitative features of the rotational velocities of stars in spiral galaxies are expected on RGT.

RGT, unlike MOND, involves no speculative hypothesis, but proceeds solely on the theoretically necessary principle of Lorentz covariance, and its origins in the problem of equal intervals of time in Newtonian physics. Therefore, even dark matter theorists, who set aside MOND, MUST take into account the special relativistic effects incorporated into RGT to estimate the amount and distribution of dark matter. *Specifically, all current estimates of dark matter and its distribution obtained by using only Newtonian gravitation are defective and unreliable, and must be recalculated using RGT.* Because RGT is a completely general theory, these remarks apply equally to the dynamics of clusters of galaxies. (To reiterate, science is NOT about “authorised knowledge”, or popularity among scientists, or “reputability”; it is about refutability.)

6 Laboratory tests of RGT

Is there any way to test RGT closer home? Indeed there is. RGT, unlike MOND, changes the gravitational force at all scales from the laboratory to the galaxy, and beyond. In a conventional Cavendish experiment, if the two attracting masses are rapidly rotated in opposite directions that would change the deflection of the suspended dumbbell, on RGT, though it would have no effect on Newtonian gravitation. In particular, if the two attracting and rotating masses are exactly lined up with the dumbbell, there would be a non-zero deflection on RGT, but zero deflection on Newtonian theory.

Of course, very high precision would be required to carry out such experiments. An unexpected difficulty here is that the value of the Newtonian gravitational constant G is not known sufficiently precisely. One reason for this is that there is an apparent discrepancy between static and dynamic ways of determining G.

According to RGT, in the dynamic way of determining G, velocity effects must be taken into account. If these are neglected, we could end up with a slightly different value of G. So, the existence of tiny discrepancies between different ways of measuring G constitutes an additional way of testing RGT. While discrepancies have indeed been noted, and it has even been speculated that these might be due to some fundamental issues,[12] the discrepancies are still within experimental error. Hopefully, these issues will be clarified in future, since it is anyway important to determine G to high precision.[13]

7 The flyby anomaly

One experiment which has already been carried out (and is likely to be repeated with greater precision) involves spacecraft when in near-earth orbit. On

RGT, one expects a tiny $\frac{v}{c}$ effect due to the rotational velocity v of the earth. The rotation of the earth has no effect on Newtonian gravitation.

Between 1990 and 2005, six NASA spacecraft flew by earth, using the technique of earth gravity assist, to either gain or lose heliocentric orbital energy. Tiny anomalies were observed [14] corresponding to an unexplained velocity difference of the order of a few mm/s at perigee. However, the observations were very precise, with systematic experimental error ranging from 0.01 mm/s to 1 mm/s, so the observations could not be put down to experimental error. Of course, the tiny anomalies may have been potentially due to many causes because the perigee velocities of the spacecraft were of the order of a few km/s, but the causes could not be explained despite a careful audit and consideration of various possible factors including general relativistic effects.[15]

Further, Anderson et al[14] found an empirical formula which fitted all six flybys:

$$\frac{\Delta V_\infty}{V_\infty} = K(\cos\delta_i - \cos\delta_o), \tag{10}$$

where ΔV_∞ was the difference between the incoming and outgoing asymptotic velocity in a geocentric frame. (Conceptually, this is the hyperbolic excess velocity at infinity of an osculating Keplerian trajectory, so the difference ought to have been zero on the Newtonian theory.) Further, δ_i and δ_o were the declinations of the incoming and outgoing asymptotic velocity vectors. The constant $K = 3.099 \times 10^{-6}$ was expressed in terms of the Earth's angular rotational velocity as ω_E (7.292115×10^{-5} rad/s), its mean radius R_E (6371 km) and the speed of light c by

$$K = \frac{2\omega_E R_E}{c}.$$

Clearly, this expression for K shows that the flyby anomaly is an effect related to the rotation of the earth, a relation expected on RGT, but inexplicable on Newtonian gravitation. If we look for a $\frac{v}{c}$ term, related to earth's rotation, using just dimensional analysis, K is the natural term that would arise. Clearly, also, if the incoming and outgoing declinations of the spacecraft are both zero (i.e., it enters and exits in the equatorial plane), then the additional RGT force which accelerates it on entry will symmetrically equal the force which retards it on exit, so there will be no net gain or loss of asymptotic velocity. A net gain or loss will arise only in the event of a difference between the asymptotic incoming and outgoing declinations. Thus, the observed anomalous effect in the flybys is *prima facie* a systematic $\frac{v}{c}$ effect depending on the rotational velocity of the earth, as expected on RGT.

Detailed modelling of the earth and exact calculations using RGT are still to be done. However, preliminary calculations already give a result which is very close. The figures below show a couple of calculations done for the Galileo (Fig. 3) and Cassini (Fig. 4) spacecraft. Past data on the orbits of these spacecraft was obtained using the NASA Horizons interface.

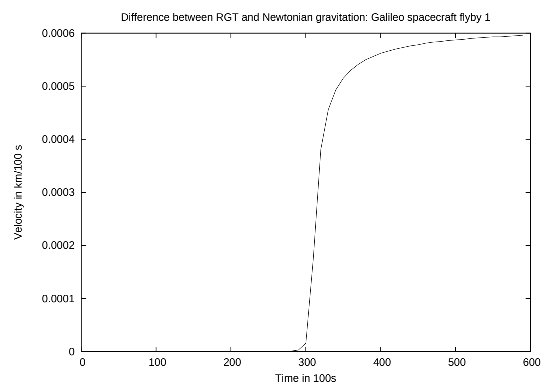


Figure 3: **Galileo.** The difference in the velocity between the solutions obtained using the new velocity-dependent RGT force and the Newtonian force, for the first flyby of the Galileo spacecraft. The x-axis is time (in units of 100 s) and the y-axis is difference of (scalar) velocity in units of km per 100 s.

The above calculations reproduce also the qualitative behaviour that most of the velocity gain or loss is close to the perigee. The computed increase or decrease in velocity is the right order of magnitude, and we expect greater accuracy with more sophisticated

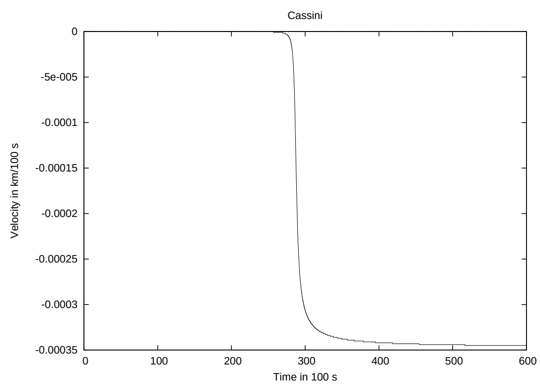


Figure 4: **Cassini.** The difference in the velocity between the solutions obtained using the new velocity-dependent RGT force and the Newtonian force for the earth flyby of the Cassini spacecraft. Same units as before. The calculated change in velocity is -3.4 mm/s compared to the reported change of -2 mm/s.

modelling of the earth.

8 Two body orbits

8.1 Laplace's argument

The Lorentz covariant RGT described above should not be confounded with the naive theories of retarded gravitation proposed over a century ago. Those naive retarded gravitation (NRG) theories were pre-relativistic and aimed to explain the discrepancy between Newtonian gravitation and the observed anomalous advance of the perihelion of Mercury. While they succeeded in that aim, they suffered from a theoretical defect: two body orbits on those theories would be unstable, as pointed out by Laplace,[16] long ago.

Thus, NRG theories typically assumed that the gravitational force pointed towards the retarded position of the attracting body (and was equal to the inverse square of the retarded distance). Laplace's objection to this was as follows. Consider two bodies in circular motion around a common centre of mass—a typical problem of Newtonian gravitation. The line of action of the NRG force would not pass through the instantaneous centre of mass. Consequently, the system would be unstable (due to a delay torque).

Laplace's argument does not apply to RGT for various reasons. First, relativistically, there is no such thing as "instantaneous centre of mass".[17] That does not mean that all relativistic theories are unstable! Further, even if one somehow defines something which can be called the instantaneous centre of mass, as some people have attempted to do, no one has proved that the "centre of mass" so defined plays the same fundamental role in deciding stability as in Newtonian mechanics. Secondly, RGT involve FDEs which do not have the same theory of stability as ODEs.

Finally, RGT differs from NRG in that the RGT force depends also upon velocity. In the above situation, of circular 2-body orbits, this means that the RGT force does NOT point directly to the retarded position of the other body, as Laplace assumed. An easy calculation shows that, in the non-relativistic case, the RGT force points closer to the instantaneous centre of mass up to $\frac{v^2}{c^2}$ terms (Fig. 5).

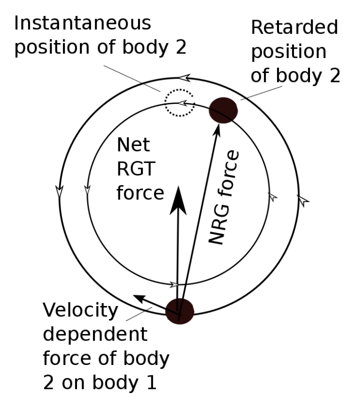


Figure 5: Difference between NRG and RGT force: Because the RGT force includes a velocity-dependent component, it points closer to the (non-relativistic) instantaneous centre of mass.

Departures from Newtonian gravitation at the $\frac{v^2}{c^2}$ level are not undesirable. Thus, for the case of Mercury, there is a long-known discrepancy with Newtonian gravitation. The classical GRT formula for the advance of perihelion ϵ, based on geodesics of the Schwarzschild solution, is

$$\epsilon = 24\pi^3 \frac{a^2}{T^2 c^2 (1 - e^2)} \tag{11}$$

where a is the semi-major axis of the ellipse, e its eccentricity, T the time period, and c the speed of light. For $e \sim 0$, when the orbit is nearly circular, $\frac{2\pi}{T}$ is an estimate of the angular velocity, and a is just the radius of the circle, so that $\frac{4\pi^2 a^2}{T^2}$ is just v^2. As such, the anomalous perihelion advance of Mercury can be regarded as approximately a $\frac{v^2}{c^2}$ effect.

Of course, since with RGT, unlike GRT, we can easily do many-body problems, the right way would be to do a many body problem, and not just linearly add up perihelia advances "due to" various causes.

8.2 Two body problem for Jupiter

We conclude with a solution of a planetary 2-body problem in RGT, as an example of how to do many body problems in RGT. The relevant equations are derived in the appendix. The equations initially involve two proper times. To solve them, we need to rewrite the equations in terms of a single coordinate time. The only non-obvious trick here is the particular 3+1 decomposition to use, as described in the appendix. (It is obvious, once we see it.)

Secondly, since these are FDEs, we need to prescribe past data. For the planetary 2-body problem, we took up the sun-Jupiter case, and prescribed past data as perfectly circular theoretical Newtonian orbits about a common centre of mass. In the Newtonian case, circular orbits remain circular, but with RGT the circle gets deformed into an ellipse, as shown in Fig. 6. However, there is no runaway instability. Had we used NRG instead, that would result in a runaway instability, as shown in Fig. 7

9 Conclusions

FDEs are an essential feature of post-relativity physics. RGT arises from modifying Newtonian grav-

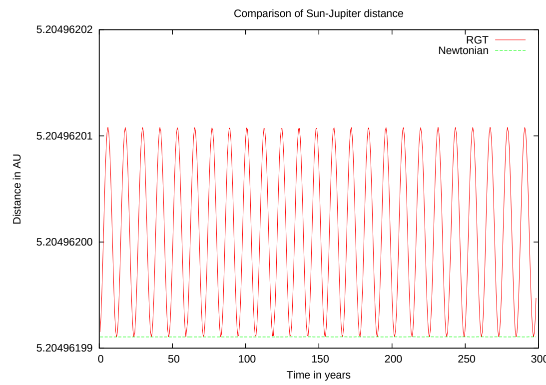


Figure 6: The relative distance of the two bodies ($\sqrt{(\vec{y}_1(t) - \vec{y}_2(t))^2}$) stays stable. The average RGT distance is marginally larger, since RGT changes the prescribed Newtonian circular orbit to an ellipse.

Figure 7: In naive retarded gravity (NRG) the force varies inversely as the square of the retarded distance and points in the direction of the retarded position of the other body. The plot brings out the instability of NRG: the distance between the two bodies ($\sqrt{(\vec{y}_1(t) - \vec{y}_2(t))^2}$) systematically increases over time in NRG.

itation to make it Lorentz covariant. This modifies the Newtonian gravitational force to include a velocity-dependent term, and the basic equations of motion of RGT are FDEs. The velocity dependent term suggests a natural explanation of the flyby anomaly as due to the rotation of the earth. Though tiny it may add up significantly across a billion stars in a spiral galaxy, which are all co-rotating. Because RGT is a theoretically essential correction to Newtonian gravitation, all galactic simulations and calculations of dark-matter must be redone using RGT. RGT does not suffer from the instability problem of pre-relativistic naive theories of retarded gravitation.

Appendix: Equations of motion in RGT

Consider two particles, with world lines given by $Y_1(s_1)$ and $Y_2(s_2)$, where Y_1 and Y_2 are 4-vectors, and s_1 and s_2 are the respective proper times. The equations of motion in RGT are

$$m_1 \frac{d^2Y_1}{ds_1^2} = F_{12}, \quad m_2 \frac{d^2Y_2}{ds_2^2} = F_{21}, \tag{12}$$

where m_1 and m_2 are the respective rest masses of the two particles, and F_{12} the 4-force exerted by particle 2 on particle 1 is given by the Lorentz covariant expression

$$\begin{aligned} F_{12} = &-\frac{kc^3}{(R_{2\,\mathrm{ret}}.V_{2\,\mathrm{ret}})^3} R_{2\,\mathrm{ret}} \\ &+ \frac{kc^3}{(R_{2\,\mathrm{ret}}.V_{2\,\mathrm{ret}})^3} \frac{(R_{2\,\mathrm{ret}}.V_1)}{(V_{2\,\mathrm{ret}}.V_1)} V_{2\,\mathrm{ret}} && (13) \\ \equiv &\left[-\frac{kc^3}{(R_2.V_2)^3} R_2 + \frac{kc^3}{(R_2.V_2)^3} \frac{(R_2.V_1)}{(V_2.V_1)} V_2 \right]_{2\,\mathrm{ret}}. && (14) \end{aligned}$$

Here, $k = Gm_1m_2$, G is the Newtonian gravitational constant, c is the speed of light, $R_{2\,\mathrm{ret}} = Y_{2\,\mathrm{ret}} - Y_1$ is the retardation vector, $V_1 = \frac{dY_1}{ds_1}$ and $V_2 = \frac{dY_2}{ds_2}$ denote the respective 4-velocities, and, in (14), $[\,]_{2\,\mathrm{ret}}$ indicates that the quantities with subscript 2 are to be evaluated at the corresponding retarded proper time, as explicitly indicated in (13). The other force F_{21} is given by interchanging 1 and 2 in (14).

In coordinates, if $Y_1 = (ct, \vec{y}_1(t))$, and $Y_2 = (ct, \vec{y}_2(t))$, the retarded coordinate time t_{12}, in the force F_{12} acting on Y_1 at time t_0, is the root of the equation

$$c^2(t - t_0)^2 = r_{12}^2 \equiv (\vec{y}_2(t) - \vec{y}_1(t_0))^2, \tag{15}$$

satisfying $t < t_0$. That is, it is the value of t at the spacetime point where the backward null cone from $Y_1(t_0)$ intersects the world line Y_2. The corresponding distance r_{12} is the retarded distance from particle 1 to particle 2. A similar equation holds for t_{21}, the retarded coordinate time in F_{21}, the asymmetry being only in the arguments of $\vec{y}_1$ and $\vec{y}_2$.

Since the two equations (12) have to be solved simultaneously, it is convenient to use a common time parameter, which we take to be the coordinate time t. We assume that the functions $t = t_1(s_1)$ and $t = t_2(s_2)$ are suitably invertible and (at least) twice continuously differentiable, and will not explicitly indicate them further. Thus, we have $\frac{dt}{ds_1} = \gamma_1$, and $\frac{dt}{ds_2} = \gamma_2$, where γ_1 and γ_2 are the respective Lorentz factors. Using an overdot to denote derivatives with respect to t, we have, by the chain rule, $V_1 = \frac{dY_1}{ds_1} = \frac{dY_1}{dt}\frac{dt}{ds_1} = \gamma_1 \dot{Y}_1$. Similarly, $\frac{dV_1}{ds_1} = \frac{dV_1}{dt}\frac{dt}{ds_1} = \gamma_1 \dot{V}_1$.

Hence, (12) can be rewritten

$$\begin{aligned} \dot{Y}_1 &= \frac{1}{\gamma_1} V_1, \\ \dot{V}_1 &= \frac{1}{\gamma_1} \frac{F_{12}}{m_1}, \end{aligned} \tag{16}$$

with similar equations for particle 2.

Since the zeroth component of these equations is not independent, we can write them in 3-vector notation using $Y_1 = (ct, \vec{y}_1(t))$, $Y_2 = (ct, \vec{y}_2(t))$, so that $\dot{Y}_1 = (c, \vec{v}_1)$, $\dot{Y}_2 = (c, \vec{v}_2)$. Let $\vec{u}_1$ and $\vec{u}_2$ denote the space components of the velocity 4-vectors V_1, and V_2, so that $\vec{u}_1 = \gamma_1 \vec{v}_1$, $\vec{u}_2 = \gamma_2 \vec{v}_2$. Further, we let $\vec{r}_{2\,\mathrm{ret}} = \vec{y}_2(t_{12}) - \vec{y}_1(t)$, denote the 3-vector corresponding to $R_{2\,\mathrm{ret}}$. Then the final equations are

$$\begin{aligned} \dot{\vec{y}}_1 &= \frac{1}{\gamma_1} \vec{u}_1 \\ \dot{\vec{u}}_1 &= \frac{1}{m_1 \gamma_1} \vec{f}_{12} \end{aligned} \tag{17}$$

where

$$\vec{f}_{12} = a\,\vec{r}_{2\,\mathrm{ret}} + b\,\vec{u}_{2\,\mathrm{ret}} \tag{18}$$

$b = a\tilde{b}$, and

$$\begin{aligned} a &= -\left[\frac{kc^3}{(R_2 \cdot V_2)^3}\right]_{2\,\mathrm{ret}} \\ \tilde{b} &= -\left[\frac{(R_2 \cdot V_1)}{(V_2 \cdot V_1)}\right]_{2\,\mathrm{ret}} \end{aligned} \tag{19a}$$

or

$$a = \left[\frac{k}{r_2^3}\right]_{2\,\mathrm{ret}} \qquad \tilde{b} = \left[\frac{r_2}{c}\right]_{2\,\mathrm{ret}} \tag{19b}$$

Here, (19b) is the non-relativistic limit of (19a).

The equations of motion (17) (accompanied by (18), (19a) or (19b)), together with the corresponding equations for particle 2 are the four 3-vector equations (or 12 equations in all) we actually solved for the sun-Jupiter problem.

References

[1] C. K. Raju. Functional differential equations. 1: A new paradigm in physics. *Physics Education (India)*, 29(3), July-Sep 2013. `http://physedu.in/uploads/publication/11/200/29.3.1FDEs-in-physics-part-1.pdf`.

[2] C. K. Raju. Functional differential equations. 2: The classical hydrogen atom. *Physics Education (India)*, 29(3), July-Sep 2013. `http://physedu.in/uploads/publication/11/201/29.3.2FDEs-in-physics-part-2.pdf`.

[3] C. K. Raju. Functional differential equations. 3: Radiative damping. *Physics Education (India)*, 30(3), July-Sep 2014. `http://www.physedu.in/uploads/publication/15/263/7.-Functional-differential-equations.pdf`.

[4] C. K. Raju. *Time: Towards a Consistent Theory*, volume 65 of *Fundamental Theories of Physics*. Kluwer Academic, Dordrecht, 1994.

[5] C. K. Raju. Time: What is it that it can be measured? *Science & Education*, 15(6):537–51, 2006. `http://ckraju.net/papers/ckr_pendu_1_paper.pdf`.

[6] C. K. Raju. *Cultural Foundations of Mathematics*. Pearson Logman, 2007.

[7] C. K. Raju. Retarded gravitation theory. In Waldyr Rodrigues Jr, Richard Kerner, Gentil O. Pires, and Carlos Pinheiro, editors, *Sixth International School on Field Theory and Gravitation*, pages 260–276, New York, 2012. American Institute of Physics. 1102:2945v3.

[8] V. C. Rubin, W. K. Ford, and N. Thonnard. Extended rotation curves of high-luminosity spiral galaxies iv. *Ap. J.*, 225:L107–L111, 1978.

[9] M. Milgrom. A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *Ap. J.*, 270:365–370, 1983.

[10] M. Milgrom. A modification of Newtonian dynamics—implications for galaxies. *Ap. J.*, 270:371–383, 1983.

[11] J. L. Synge and B. A. Griffith. *Principles of Mechanics*. McGraw Hill, 3rd edition, 1959. Equation 18.320.

[12] Stephen Schlamminger, Eugene Holzschuh, Walter Kündig, Frithjof Nolting, and Jürgen Schurr. Determination of the gravitational constant. In C. Lämmerzahl, C. W. F. Everitt, and F. W. Hehl, editors, *Gyros, Clocks, Interferometers...: Testing Relativistic Gravity in Space*, pages 15–28. Springer, Berlin, 2001.

[13] Jun Luo, Qi Liu, Liang-Cheng Tu, Cheng-Gang Shao, Lin-Xia Liu, Shan-Qing Yang, Qing Li, and Ya-Ting Zhang. Determination of the Newtonian gravitational constant G with time-of-swing method. *Physical Review Letters*, 102(24):240801(4), June 2009.

[14] John D. Anderson and et al. Anomalous orbital-energy changes observed during spacecraft flybys of earth. *Physical Review Letters*, 100:091102, 2008.

[15] C. Lammerzahl, O. Preuss, and H. Dittus. Is the physics within the solar system really understood? 2006. arXiv.gr-qc/0604052.

[16] P. S. Laplace. *Celestial Mechanics*, volume 4, Book 10, chp. 7. Little and Brown, Boston, 1839. trans. N. Bowditch.

[17] M. H. L. Pryce. The mass centre in the restricted theory of relativity.... *Proc. R. Soc. A*, 195:62–81, 1948.

Creation of electric potential minima in 3D: Trapping charged particles

P. Mandal

Department of Physics
St. Paul's Cathedral Mission College
33/1 Raja Rammohan Roy Sarani, Kolkata 700009, India.
pintuphys@gmail.com

(Submitted 28-05-2015)

Abstract

An electrostatic potential minima does not exist in three dimension. However, an alternating electric field can produce a dynamical potential minima in three dimensional space and charged particles can be trapped within such potential well. A system of trapped ion/s is almost free from unknown external perturbations and hence such a system finds enormous applications in different fields. This article explains how such a potential minima can be developed with electric field only and how the charged particles can be trapped within it. Some important applications of trapped particles have been outlined here with a demonstrative experiment for realization of the technique.

1 Introduction

'Let us consider a particle at rest'-this is often the introductory sentence in our text books or while teaching in classroom. But can we have a particle at rest in practice? A famous remark from Erwin Schroedinger may be quoted in this regard, 'We never experiment with just one electron or atom or (small) molecule. In thought-experiments we sometimes assume that we do; this invariably entails ridiculous consequences'. However, it has now become a reality to have a particle almost at rest. A single particle like an atom, or even an electron can be confined in space within a region of few micrometers. For confinement of charged particles, two different techniques were developed by two pioneers Wolfgang Paul (the device, named after him, known as Paul trap [1]) and Hans Georg Dehmelt (the device, named after Frans Michel Penning, known as the Penning trap [2]). In Paul trap, the charged particles can be trapped by using a static electric field together with a time varying electric field while in Penning trap, a static magnetic field is required in association with a static electric field. Both of these devices are regularly used as important tool in different fields, both of fundamental physics interests and commercial applications. Here we will restrict ourselves in discussions related to the Paul trap. The readers are, however, referred to an article [3] which covers discussions on both the Paul trap and the Penning trap.

The article has been arranged in the following way. In section 2, the fundamental technique of creation of potential minima using only the electric field (as associated with the Paul trap) has been described. The equation of motion of a single charged particle within such a dynamic potential well is reviewed in section 3. In section 4, a demonstrative experiment has been presented. The applications of trapped ion system in different fields have been outlined at the end of this article (section 5).

2 How to create a potential minima in 3D?

A particle in one dimension can be confined by a restoring force proportional to its displacement from the equilibrium position (the force as associated with a simple harmonic motion). In other words, it requires a quadrupole potential (proportional to the square of the displacement). Naturally, for three dimensional trapping of a particle, the potential should be quadrupolar in all three dimensions and is described as follows:

$$\Phi(x,y,z) = Ax^2 + By^2 + Cz^2, \qquad (1)$$

where A, B, C are constants. For a charged particle, this potential can be chosen as the electric potential. Thus the force on a particle of charge e under the influence of this potential is given by

$$\begin{aligned} \vec{F}(x,y,z) &= -e\vec{\nabla}\Phi(x,y,z) \\ &= -2e(Ax\hat{x} + By\hat{y} + Cz\hat{z}). \qquad (2) \end{aligned}$$

As is necessary for trapping, the force $\vec{F}$ should be restoring in nature and thus it follows that the constants A, B and C are all positive (for a positively charged particle). However, any electrostatic potential in free space should satisfy the Laplace's equation ($\nabla^2\Phi(x,y,z) = 0$), following which, at least one constant must be negative for this electrostatic potential. Thus it can be concluded that no electrostatic potential minima exists in three dimension [1].

So how to create the electric potential minima in three dimension? The answer to the this challenge was addressed by Wolfgang Paul who demonstrated that a time varying electric potential can produce a 'dynamic minima' in three dimension. The idea is to vary A, B and C with respect to time, such that the potential having its minima in one direction at an instant, rotates to the other direction at a later instant. If the rotation of the potential minima is faster as compared to the motion of the charged particle, the particle will experience a time-averaged potential minima in all directions. The particle will be confined within the potential well if the average potential depth is larger than its kinetic energy. This can be compared to a ball placed on a rotating saddle (see, for reference, a nice demonstration in youtube, the mechanical analogue of Paul trap [4]).

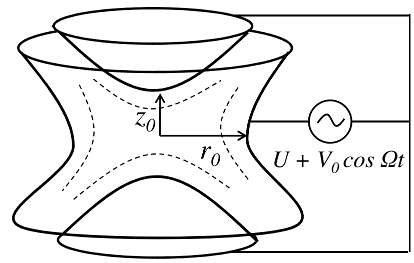


Figure 1: The hyperbolic geometry of the electrodes with necessary electrical connections for developing the quadrupole potential as defined in eqn. 5. The dotted lines show hyperbolic equipotential surfaces (eqn. 3).

In order to produce the quadrupole electric potential, suitable geometry of the electrodes is required. If there exists a rotational symmetry about the z axis, $A = B$, and hence $C = -2A$. Consider, for example, the electrode geometry depicted in fig. 1. As can be seen from fig. 1, two coaxial bowl-shaped electrodes at the ends, together with the ring electrode at the middle are hyperboloids of revolution about the z axis. If the radial and axial dimensions of the trap are respectively r_0 and z_0, the equations for the hyperbolic electrode surfaces are given by [5]

[1] This is, in literature, known as Earnshaw's theorem

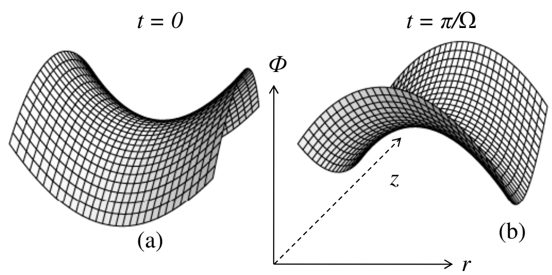


Figure 2: Potential surface in the $r-z$ plane at different instant. (a) Harmonic oscillator Potential along r(in $x-y$ plane) and inverted oscillator potential along z at $t = 0$. (b) Harmonic oscillator Potential along z and inverted oscillator potential along r at $t = \pi/\Omega$. The minima of the potential surface rotates in the $r-z$ plane with angular frequency Ω, the frequency of the applied alternating potential.

$$\begin{aligned} r^2 - 2z^2 &= r_0^2, \\ r^2 - 2z^2 &= -2z_0^2, \end{aligned} \tag{3}$$

where $r^2 = x^2 + y^2$. When voltage is applied to the middle electrode or on the end electrodes, it produces equipotential surfaces defined by eqn. 3. The potential inside the trap can therefore be written as

$$\Phi(r, z) = A(r^2 - 2z^2). \tag{4}$$

Now, the coefficient A should be chosen in such a way that the potential has its minima along r (i.e. in the $x-y$ plane) at an instant, and in the z direction at a later instant. To elucidate the statement, let us consider the following form of the potential:

$$\Phi(r, z, t) = \frac{U + V_0 \cos \Omega t}{2r_0^2}(r^2 - 2z^2). \tag{5}$$

It is seen from eqn. 5 that, at time $t = 0$ the potential resembles that of simple harmonic oscillator along r and inverted harmonic oscillator along z [fig. 2(a)]. However, the vice-versa hold at $t = \pi/\Omega$ [fig. 2(b)].

3 Motion of a Trapped Ion

The equation of motion of a single particle of charge e and mass m under the influence of the potential (defined by eqn. 5) follows from Newton's law of motion and can be described by the following equations:

$$\begin{aligned} \frac{d^2r}{dt^2} &= -\frac{e}{mr_0^2}(U + V_0 \cos \Omega t)r, \\ \frac{d^2z}{dt^2} &= \frac{2e}{mr_0^2}(U + V_0 \cos \Omega t)z. \end{aligned} \tag{6}$$

Using a common symbol u for both r and z, and introducing the dimensionless parameters a_u, q_u and ζ the eqn. 6 can be rewritten as

$$\frac{d^2u}{d\zeta^2} + (a_u - 2q_u \cos 2\zeta)u = 0, \tag{7}$$

where

$$\begin{aligned} a_z &= -2a_r = -\frac{8eU}{mr_0^2\Omega^2}, \\ q_z &= -2q_r = \frac{4eV_0}{mr_0^2\Omega^2}, \\ \zeta &= \frac{\Omega t}{2}. \end{aligned} \tag{8}$$

The equation of motion (eqn. 7) is a standard differential equation in mathematics, known as Mathieu differential equation. The solutions of this equation result in either stable or unstable motion depending on the values of the parameters a_u and q_u, defined in eqn. 8. There exists a region in a_u vs. q_u diagram for which the ion-motion is stable along a particular direction, for example along r (fig. 3). A similar stability region exists for the motion along z direction. An intersection between these two stability regions (the shaded region in fig. 3) is where the stable motion in three dimension is sustained.

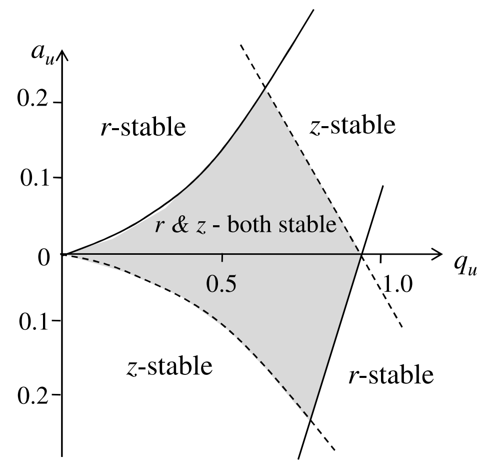


Figure 3: Stability region in a_u vs. q_u diagram for an ion trap. The regions bounded by the dotted line and bold line correspond to stable motion along r and z respectively. The motion is stable along both r and z directions in the shaded region and the trap operating parameters a_u and q_u are chosen in this region.

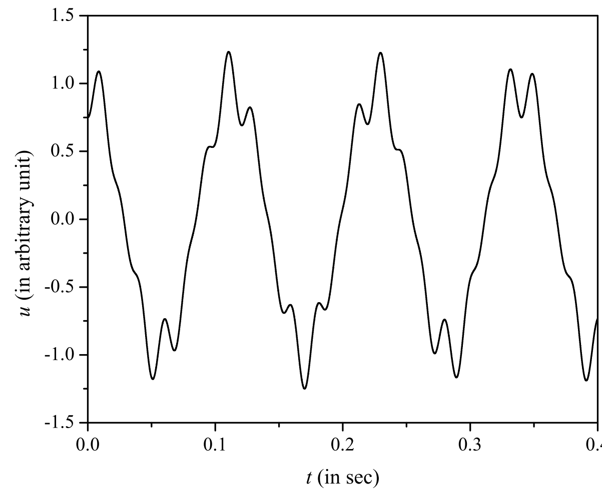


Figure 4: The motion of a trapped ion in one direction. A low frequency motion, called the secular motion is modulated by a high frequency motion, the micromotion. The simulation of the trajectory is done for $q_r = 0.5$, $\Omega = 2\pi \times 50$ rad/s and $\omega_{0r} = 2\pi \times 8.5$ rad/s.

In the 'adiabatic approximation' (i.e. for small a_u and q_u values), the solution of Mathieu differential equation can be represented in the following form [5]:

$$u = c\left(1 - \frac{q_u}{2}\cos\Omega t\right)\cos\omega_{0u}t, \tag{9}$$

where c is a constant and

$$\omega_{0u} = \frac{\beta_u \Omega}{2}. \tag{10}$$

The parameter β_u, for small a_u and q_u, can be defined as

$$\beta_u \approx \sqrt{a_u + \frac{q_u^2}{2}}. \tag{11}$$

Eqn. 9 shows that the ion oscillates with a frequency ω_{0u} and its motion is modulated with the frequency Ω of the applied alternating potential (fig.4). For $\beta_u < 1$, $\omega_{0u} < \Omega$. The slow frequency motion (at ω_{0u}) is called the macromotion or secular motion while the higher frequency motion (at Ω) is termed as the micromotion.

4 An Experiment

In this section, a demonstrative experiment has been described. Dust particles, here chalk dust, have been trapped in a ring trap at the line frequency, at 50 Hz.

The trap setup is shown schematically in fig. 5. The surfaces of two end electrodes are hyperboloids of revolution about the z axis and the ring electrode at the middle has hyperbolic cross section, a similar geometry that is described in fig. 1. In the experiment, the electrodes are made of brass. The ring is taken of diameter ~ 10 mm ($r_0 = 5$ mm) and the end cap electrodes are separated by a distance ($2z_0$) of 7 mm (note that, $r_0^2 = 2z_0^2$, a dimensional constraint

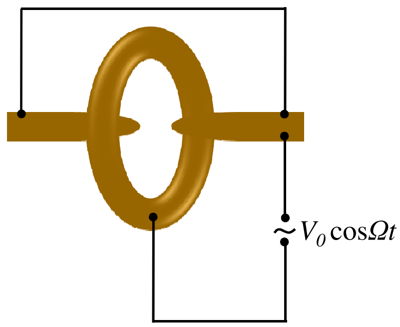


Figure 5: Schematic of a ring trap setup. The end electrodes are connected together and the alternating potential is applied between the ring and the end electrodes.

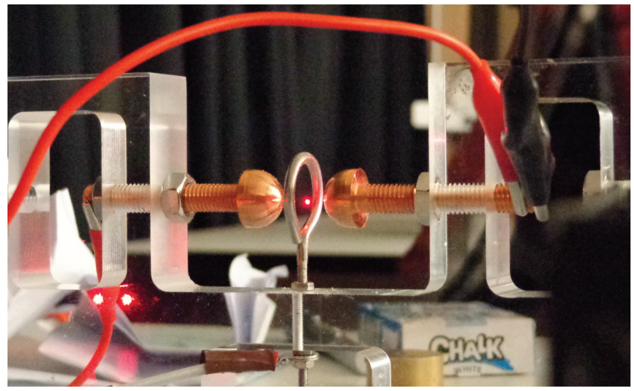


Figure 6: A photograph of the setup with trapped chalk dust. A laser beam is incident on the trapped dust at the center and is scattered by the dust particles for clear visualization of ion trapping.

of this electrode geometry as necessary for efficient trapping). The end electrodes are electrically connected together and an alternating voltage is applied between the ring electrode and the end electrodes. The line voltage (230 V, 50 Hz) is passed through a variac and fed to a step-up transformer for necessary voltage amplification. A typical voltage used for trapping is 1500 V. It is to be noted that no dc potential is applied here i.e. $U = 0$ and hence $a_u = 0$ (dc potential just modifies the effective potential depth). The chalk dust are taken in a syringe and injected inside the trap. The dust get ionized due to injection and are trapped inside. A photograph of the experimental setup with trapped dust particles at the center is presented in fig. 6.

The dust particles form thread-like clusters and oscillate inside the trap. If the trapping voltage is stabilized and the system is adequately isolated from the surroundings, the particles can be stored for days within the trap. It is possible to estimate the charge-to-mass ratio of the trapped dust clusters. For stable and efficient trapping, the q parameter should be around 0.5. With the applied ac voltage $V_0 = 1800$ V, at frequency $\Omega = 2\pi \times 50$ rad/s, the charge-to-mass ratio (e/m) is estimated from eqn. 8 as 3×10^{-4} C/kg.

5 Applications

Ion traps provide best realization of 'particle at rest' and hence it is used as an important tool in many applications. The thermal motion of trapped ion can be reduced by laser cooling technique and it can be localized within its de Broglie wavelength which is few μm [6]. Thus a single trapped and laser cooled ion represents a perturbation-free quantum system. A series of experiments are being performed for testing and demonstrating wide aspects of fundamental physics. A single or few ions are used for precision measurement of various atomic properties such as lifetime of atomic states [7], transition frequency or ac Stark shift [8], quadrupole moment of atomic states [9], atomic parity violation [10] *etc.* A single trapped ion is used for developing atomic frequency standard [11]. Single or few trapped ions are used for quantum teleportation [12], quantum information processing [13] and designing quantum computer [14]. Large ion traps are used for Coulomb crystal study [15], mass spectrometric applications [16] and many more [17].

6 Acknowledgments

The author thank all those who helped in this demonstrative experiment at the Indian Association for the Cultivation of Science, Kolkata. The photograph of the experiment is due to courtesy of S. Das, National University of Singapore, Singapore.

References

[1] W. Paul (1990). Rev. Mod. Phys., 62, 531

[2] H. G. Dehmelt (1990). Rev. Mod. Phys., 62, 525

[3] P. M. Rao, R. D'Souza, and S. A. Ahmad, (2001). Resonance, 6(04), 22

[4] www.youtube.com/watch?v=XTJznUkAmIY

[5] F. G. Major, V. N. Gheorghe, and G. Werth Charged Particle Traps (Part I), (Springer-Verlag, Heidelberg, 2005)

[6] D. Leibfried *et al.*, (2003). Rev. Mod. Phys., 75, 281

[7] N. Yu, W. Nagourney, and H. Dehmelt (1997). Phys. Rev. Lett., 78, 4898

[8] J. A. Sherman *et al.*, (2005). Phys. Rev. Lett., 94, 243001

[9] W. H. Oskay, W.M. Itano, and J. C. Bergquist, (2005). Phys. Rev. Lett., 94, 163001

[10] N. Fortson (1993). Phys. Rev. Lett., 70, 2383

[11] R. J. Rafac *et al.*, (2000). Phys. Rev. Lett., 85, 2462

[12] M. Hotta (2009). Phys. Rev. A, 80, 042323

[13] J. Benhelm, G. Kirchmair, C. F. Roos, and R. Blatt (2002). Phys. Rev. A, 77, 062306

[14] D. Kielpinski *et al.*, (2002). Nature, 417, 709

[15] M. Drewsen *et al.*, (1998). Phys. Rev. Lett., 81, 2878

[16] R. E. March, and J. F. J. Todd Quadrupole Ion Trap Mass Spectrometry (second edition), (Wiley, New York, 2005)

[17] G. Werth, V. N. Gheorghe, and F. G. Major Charged Particle Traps II (Applications), (Springer-Verlag, Heidelberg, 2009)

Note on the magnetic energy of a rotating charged metal sphere

Sergey N. Sazonov[1]

[1]Ufa State Aviation Technical University
Ufa 450000, Bashkortstan Republic, Russia.

(*Submitted* 12-01-2015)

Abstract
Contribution of the magnetic energy to the energy of field of a rotated charged conducting spheroid is calculated by a method which does not use either the integration of the magnetic energy density as the surface integration of the scalar product of the current density vector or the vector potential of field. The method may be interesting for a student, studying the classical electrodynamics course.

1. Introduction

The problem of magnetic field calculation, created by a rotating charged conducting sphere, is a traditional part of university textbooks on electrodynamics ([1]). If the sphere's radius is a, it's angular velocity is ω and the net charge on sphere is $Q = \sigma_0 \cdot 4\pi a^2$, then, in Gaussian units, the solution of the Poisson equation for the vector potential $\boldsymbol{A}(\boldsymbol{R})$

$$\nabla^2 \boldsymbol{A} = -\frac{4\pi}{c}\boldsymbol{j}, \qquad (1)$$

with the current density

$$\boldsymbol{j}(\boldsymbol{R}) = \sigma_0 \cdot \omega a \sin\theta \cdot \delta(R-a)\boldsymbol{e}_\varphi \qquad (2)$$

is

$$\boldsymbol{A}(r<a) = \frac{1}{2} B \cdot r \sin\theta \cdot \boldsymbol{e}_\varphi, \qquad (3)$$

$$\boldsymbol{A}(r \ge a) = \frac{\Re}{r^2} \cdot \sin\theta \cdot \boldsymbol{e}_\varphi. \qquad (4)$$

Here

$$\Re = \frac{Qa^2}{3c}\boldsymbol{\omega}, \qquad (5)$$

is the sphere's dipolar moment,

$$\boldsymbol{B} = \frac{2}{a^3}\Re, \qquad (6)$$

is the magnetic field induction inside it, R, θ, φ are the spherical coordinates, defined so that Oz axis is along the vector $\boldsymbol{\omega}$ and the center of sphere corresponds to $\boldsymbol{R} = \boldsymbol{0}$, $\boldsymbol{e}_\varphi$ is the correspondent unit vector.

To tell strictly, even in the zero electron mass approximation, which will be used further, the action of the Lorenz force on electrons will disturb the surface charge density σ_0 (we can consider the thin metal film sputtered on a dielectric ball instead of to consider the metal ball). However, it is easy to show (as in [2]), that the relatively rearrangement $\Delta\sigma/\sigma_0$ of the charge density at any point of sphere is proportional to β^2 where $\beta = \omega a / c$ (c is the speed of light in vacuum). Considering $\beta << 1$ we will not pay attention on this effect beneath.

To spin the ball with charged metal film, it's necessary to spend the work W against the eddy electric field among other. The quantity W is called the magnetic field energy. Contrary to (3) – (6), the W value is not presented in [1]. Meanwhile this quantity is called for the theoretical physics as it is seen from the original journal articles. In [3]

(Appendix), author uses for the calculation of W the formula

$$W = \frac{1}{8\pi}\int_{R^3} \boldsymbol{B}_{\mathrm{rot}}^2 dV, \qquad (7)$$

(R^3 is the symbol of integrating over all space, $\boldsymbol{B}_{\mathrm{rot}}(\boldsymbol{R})$ is the magnetic field induction in an arbitrary point) and derives

$$W = \frac{\Re^2}{a^3}. \qquad (8)$$

In [4] (Appendix) authors computed (8) with the help of formula

$$W = \frac{1}{2c}\int_{R^3} \boldsymbol{A}(\boldsymbol{R})\cdot \boldsymbol{j}(\boldsymbol{R}) dV, \qquad (9)$$

what is the more simple way owing to Dirac function presence in (2). The aim of this note is to show that an undergraduate student studying the classical electrodynamics may not spend the time for reproducing the routine algebraical calculations and to derive (8) more simply than in [4] if he orients freely in the theme of magnetostatics of ferromagnets in the volume, for example, of [5]. The method of calculation of the rotating charged body magnetic energy will be applied to the spheroid. For auditory purposes an educator may adapt this method turning the spheroid to sphere primarily and expelling a part of mathematics beneath.

2.Magnetic energy of a rotating charged metal spheroid

Let the thin metal film is sputtered onto the dielectric so that the equation of the external metal surface is:

$$\left(\frac{\rho}{a}\right)^2 + \left(\frac{z}{b}\right)^2 = 1, \qquad (10)$$

where $\rho = \sqrt{x^2 + y^2}$. The distribution of the surface charge density $\sigma(z)$ is ([6]):

$$\sigma(z) = \frac{Q}{4\pi b}\frac{1}{\sqrt{\rho^2 + \left(\frac{a}{b}\right)^4 z^2}}. \qquad (11)$$

After the body began to rotate the linear current density of the surface charge is

$$i(z) = \sigma(z)\cdot\omega\rho(z) = \frac{Q\omega}{4\pi b}\cdot\cos\alpha(z), \qquad (12)$$

where the angle α is defined in the Figure 1 and the equation $\mathrm{tg}\alpha = d\rho / d|z| = z \cdot a^2 / \rho \cdot b^2$ for the spheroid surface was used.

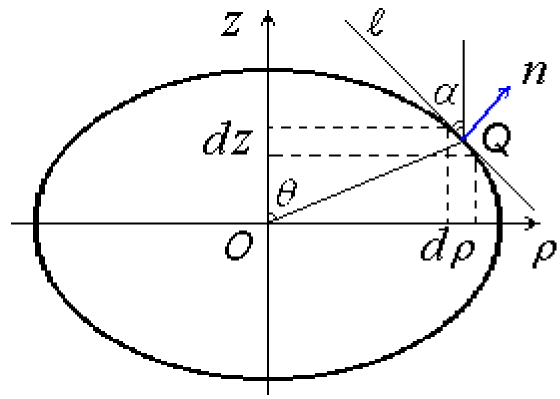


FIG. 1: Vertical section of metal spheroid (the dielectrical core is not shown). *Q* – arbitrary point on it's surface. *n* – normal to surface in *Q*, ℓ is the line, tangent to spheroid in Q.

The magnetic field induction $\boldsymbol{B}_{\mathrm{rot}}(\boldsymbol{R})$ created in all space by such current distribution coincide with the just one $\boldsymbol{B}_{\mathrm{magn}}(\boldsymbol{R})$ created by the spheroidal permanent magnet of the same shape (10) with the magnetization

$$\boldsymbol{M} = \frac{1}{c}\cdot\frac{Q\omega}{4\pi b}\boldsymbol{e}_z, \qquad (13)$$

in accordance with the well-known statement ([6]) that the uniform magnetization $\boldsymbol{M}$ is equivalent to molecular current with the linear surface density

$$\boldsymbol{i} = c\cdot[\boldsymbol{M};\boldsymbol{n}], \qquad (14)$$

where $\boldsymbol{n}$ is the unit vector of outer normal at the given point of magnet surface. Let us call such a magnet as the equivalent one for our rotating charged metal spheroid.

Let the two ideal conductors, screaning completely any external variable magnetic field inside it at zero temperature, move from infinity to the equivalent magnet as it is shown in Figure 2. Everyone of the two has one flat surface with the deepening as the half of our spheroid. When the conductors will taught the magnet, the field of screaning currents $\boldsymbol{B}_{\mathrm{j}}(\boldsymbol{R})$ will compensate the field of magnet in all space so that $\boldsymbol{B}_{\mathrm{j}}(\boldsymbol{R}) + \boldsymbol{B}_{\mathrm{magn}}(\boldsymbol{R}) = 0$. So the magnetic energy (7) of our rotating spheroid is equal to the magnetic energy of currents in conductors which is, by definition, the work of external forces $\boldsymbol{F}_1$, $\boldsymbol{F}_2$ being spended for the conductors transition from

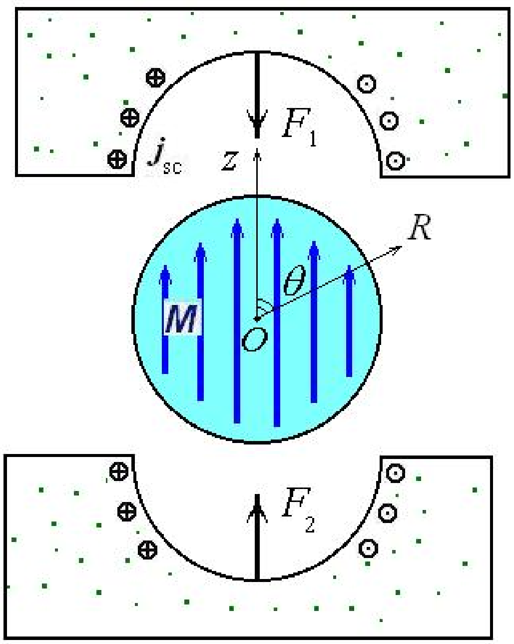


FIG. 2: Permanent magnet and the ideal conductors. $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are the external forces transferring the last two. ***J***sc are the screening currents appeared by this.

infinity. This work will increase the free energy $\Im$ of the system "magnet + ideal conductors":

$$\Im = W + \Im_{\mathrm{ext}}\,, \tag{15}$$

where $\Im_{\mathrm{ext}}$ is the free magnetostatic energy of the retired equivalent magnet. As it was discussed in [7],

$$\Im_{\mathrm{ext}} = -\frac{1}{2}\boldsymbol{M}\boldsymbol{B}'\cdot V = -\frac{1}{2}\boldsymbol{M}\boldsymbol{H}\cdot V + k\boldsymbol{M}^2\,, \tag{16}$$

where $\boldsymbol{B}'$ is the micro field, acting on the magnetic moments inside the spheroid, $\boldsymbol{H}$ is the magnetic field strength there, k depends only of a crystal lattice type of material of magnet and V is it's volume. (k is ignored often in literature as it is made in [5]) Formula for $\Im$ is presented in [8]:

$$\Im = -\frac{1}{2}\boldsymbol{M}\left(\boldsymbol{B}' + \boldsymbol{B}_{\mathrm{j}}\right)\cdot V\,, \tag{17}$$

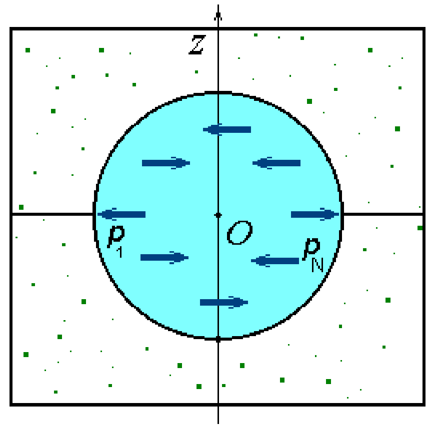


FIG. 3: The demagnetized magnet inside the system of two ideal conductors. $\boldsymbol{P}_1$ – $\boldsymbol{P}_N$ are the magnetic moments of atoms.

without any comments and references what reflects it's evidence for professional physics-theoreticians. For students the next comment may be done. Let us imagine that the equivalent magnet is magnetized in the presence of the ideal conductors enveloping it from the initial state where the magnetic moments $\boldsymbol{p}_i$ of atoms (i = 1 … N) lie chaotically in the Oxy plane (Figure 3). During the magnetization process $\boldsymbol{p}_i$ vectors rotate up to Oz axis so that the value of angle θ between $\boldsymbol{p}_i$ and Oz axis is just the same for all atoms at any moment of time ($\theta = \pi / 2$ in Figure 3). Then the vectors $\boldsymbol{B}'$ and $\boldsymbol{B}_{\mathrm{j}}$ became functions of θ and ($\boldsymbol{d\theta}_i$ is the vector of infenitizemal rotation of $\boldsymbol{p}_i$):

$$\Im = N\cdot\int_{\theta=\pi/2}^{\theta=0}\left[\boldsymbol{p}_i;\boldsymbol{B}'(\theta) + \boldsymbol{B}_{\mathrm{j}}(\theta)\right]\cdot\boldsymbol{d\theta}_i =$$

$$= N\cdot\int_{\theta=\pi/2}^{\theta=0}\left(\boldsymbol{B}'(\theta) + \boldsymbol{B}_{\mathrm{j}}(\theta)\right)\cdot\left[\boldsymbol{d\theta}_i;\boldsymbol{p}_i\right]. \tag{18}$$

Taking in (18) $\boldsymbol{B}_{\mathrm{j}} = 0$ and $\boldsymbol{B}'(\theta) \sim \cos\theta \cdot \boldsymbol{e}_{\mathrm{z}}$, we derive (16). But for the screaning currents it must be $B_{\mathrm{j}}(\theta) \sim B'(\theta)$ as the consequence of the Maxwell equation rot $\boldsymbol{H} = 4\pi \boldsymbol{j} / c$ linearity. So at any θ, it will be $\boldsymbol{B}_{\mathrm{j}}(\theta) \sim \cos\theta \cdot \boldsymbol{e}_{\mathrm{z}}$ and we came to (17).

Using (16) – (17) in (15) we obtain

$$W = \frac{1}{2}\boldsymbol{B}_{\mathrm{magn}} \cdot \mathfrak{R} = \frac{1}{2}\left(4\pi\boldsymbol{M} + \boldsymbol{H}\right)\cdot\mathfrak{R} =$$

$$= \mathfrak{R}\cdot\frac{1}{2}\left(4\pi\boldsymbol{M} - N_z\boldsymbol{M}\right) = N_x \cdot \frac{\mathfrak{R}^2}{V}, \qquad (19)$$

where N_x and N_{z} are the demagnetization factors of spheroid along Ox and Oz axes and, as it resulted from (13),

$$\mathfrak{R} = \boldsymbol{M}\cdot V = \boldsymbol{M}\cdot\frac{4}{3}\pi a^2 b = \frac{Qa^2}{3c}\boldsymbol{\omega}. \qquad (20)$$

For the sphere $N_x = 4\pi / 3$ and we return from (19) to (8). In the general case of the arbitrary $m = b / a$ it is conveniently to present the result as the W / W_{el} dependence of m where $W_{\mathrm{el}} = Q^2 / 2C$ is the electrostatic energy of the charged metal spheroid and C is it's capacity. Taking the expressions for C from [6] and for N_x from [9] we have:

$$\frac{W}{W_{\mathrm{el}}} = \frac{\beta^2}{3}\cdot\frac{1}{1-m^2}\left\{1 - \frac{m\cdot\sqrt{1-m^2}}{\arccos m}\right\} + O(\beta^4), \quad (21)$$

for the oblate spheroid and

$$\frac{W}{W_{\mathrm{el}}} = \frac{\beta^2}{3}\cdot\frac{1}{m^2-1}\left\{\frac{m\sqrt{m^2-1} + \frac{1}{2}\cdot\ln\left(m - \sqrt{m^2-1}\right)}{\ln\left(m+\sqrt{m^2-1}\right)} - \frac{1}{2}\right\} + O(\beta^4). \qquad (22)$$

for the prolate one. Formulae (21) – (22) are illustrated graphically in the Figure 4.

Formula (19) with $\mathfrak{R}$ derived from (20) must became the strict one for infinite cylinder when the tangential component of the Lorenz force acting on the free electrons disappears. Inserting $\boldsymbol{B} = 4\pi\boldsymbol{M}$ in (20) and using (20) in (19) with $N_x = 2\pi$, we return to (7).

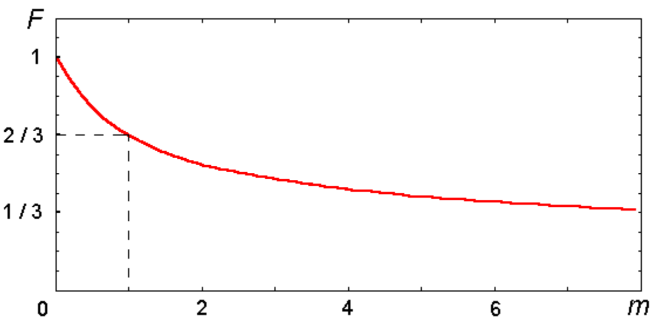


FIG. 4: The dependence of the value $F = (3 / \beta^2)\cdot W / W_{\mathrm{el}}$ as function of the spheroid parameter $m = b / a$

References :

[1] V. V. Batygin, I. N. Toptygin, *Problems in Electrodynamics*, 2nd ed. (Academic, London, 1978).
[2] K.T. McDonald "*Charged, conducting, rotating sphere*", http://www.physics.princeton.edu/~mcdonald/examples/
[3] D. Lynden-Bell, "*The Relativistically Spinning Charged Sphere*", Phys. Rev. D **70**, 104021-1–104021-14 (2004).
[4] M. C. N. Fiolhais, H. Essen, C. Providencia, A. B. Nordmark, "*Magnetic field and current are zero inside ideal conductors*", Progress In Electromagnetics Research B **27**, 187–212 (2011).
[5] S. Chikazumi, *Physics of Ferromagnetism*, 2nd ed. (Clarendon Press, Oxford, 1997), ch. 1.
[6] J.A. Stratton, *Electromagnetic theoria* (McGraw – Hill, New York, 1941).
[7] W. F. Brown, *Micromagnetics* (Interscience, New York, 1963).
[8] Yu. I. Bespyatkh, V. Vasilevski, E.G. Lokk, V. D. Kharitonov , "*Suppression of the domain structure in uniaxial ferromagnetic films with a superconducting coating*", Phys. of the Solid State **40**, 975 – 981 (1998)
[9] J. A. Osborn, "*Demagnetizing factors of the general ellipsoid*", Ph. Rev. **67**, 351 – 357 (1945).

The quark confinement: Life sentence of fundamental constituents of nature

Haresh Raval

Department of Physics
Indian Institute of Technology,
Bombay, Mumbai - 400076, India
haresh@phy.iitb.ac.in

(Submitted 28-05-2015)

Abstract

The particle physics is a study of the properties and interactions of subatomic particles. Here we discuss the basics of one of the most crucial phenomena in particle physics, namely, The quark confinement. Subsequently, we describe The Bag Model of the quark confinement, which makes a phenomenology of the confinement easy to understand.

1 Introduction

The idea of quarks was put forward by Murray Gell-Mann and George Zweig in the year 1964 [1]. Later, experiments established the quarks as real physical, fundamental objects. There are six different types of the quark, namely, Up, Down, Top, Bottom, Charm and Strange. They all have an electric charge. Down, Bottom and Strange quarks have an identical charge of $-\frac{1}{3}e$; whereas Top, Up and Charm quarks have an identical charge of $\frac{2}{3}e$. All of them are massive with the top quark being the heaviest and the up quark being the lightest. Up and Down quarks are the only stable quarks. This is the reason why the ordinary matter, we see around us, is made up of them. The rest are created in high energy collisions only and quickly decay into these quarks. Each quark has its antiparticle called antiquark. A subatomic particle made of quarks is commonly known as 'Hadron'. There are two types of the hadron, namely, Meson and Baryon. Meson is composed of the quark and the antiquark of different types e.g. Pion$^+$ is composed of the up and the anti-down quark pair. Baryon is composed of three quarks e.g. Neutron is composed of two down quarks and one up quark. Having given a brief introduction about the quarks and the recipe to understand the confinement, we now proceed to the main subject of this article.

2 The quark confinement

Experimental developments over many years have confirmed the non-existence of free quarks unlike other fundamental particles such as the electron and the neutrino. What is observed instead in experiments is jets of hadrons. For example, when protons (a known member of the baryon family) or nuclei of heavy atoms such as lead are smashed in particle colliders like large hadron collider, products of collisions are jets of mesons and baryons. One such collision is shown below

$$proton + proton \rightarrow proton + neutron + pion^{+}$$

These experiments have unearthed the truth, which is against our day-to-day experience in a following sense. When we break something assembled, it shatters into more basic objects which make that thing and not into other similar entities. For example, when a building is blasted, it collapses into debris of cement concrete, bricks and other materials which made it and not into other different types of buildings. However, what happens after collision of protons or nuclei is exactly opposite. Parent hadrons are not broken into quarks but get converted into different daughter hadrons after collisions. Thus, the quarks by default clump together to form hadrons. So, the quarks are forever trapped inside hadrons ('life sentence'), a phenomenon known as the confinement.

The physics of the confinement phenomenon is 'the asymptotic freedom', a peculiar property of the strong force through which quarks interact. Naively, the asymptotic freedom means that the quarks do not interact with one another when they are very close and as the separation increases, strength of interaction keeps growing. Therefore, quarks tend to be close together. Again, this property contradicts a familiar classical physics. When two charges or masses are brought closer to each other, the electromagnetic or the gravitational force between them increases as $\frac{1}{r^2}$, inverse square of the distance between them, but the strong nuclear force has an opposite behavior. The phenomenology of the confinement physics can be easily understood by a simple model called "The Bag Model".

3 The Bag Model

The Bag Model was developed in 1974 by a group of physicists at the Massachusetts Institute of Technology in Cambridge (USA) [2] and soon it became popular in particle physics community. In this model, mesons are considered as elastic bags containing quark-antiquark pair and baryons are considered as elastic bags containing three quarks. When any elastic material undergoes deformation due to external force, it tries to resist the deformation. The same is true for these imaginary elastic bags but they are little weird since when the original bag is broken, new bags are formed out of it automatically.

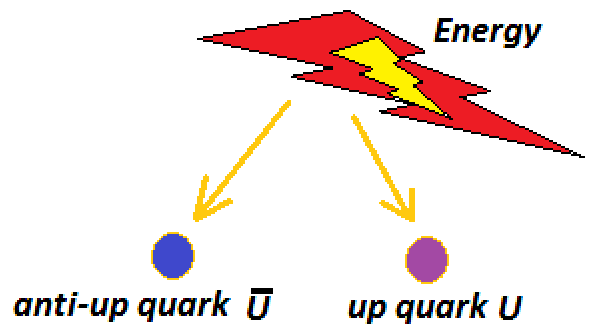


Figure 1: Example: Energy being converted into up-antiup quark pair

Imagine an elastic bag as shown in fig. 2 containing quarks which represents a baryon. When a bag is unstretched, quarks can move freely in it. As soon as you try to pull a quark apart, the bag stretches and resists. More you stretch, more the energy required to stretch it even further. In other words, potential energy of two quarks is proportional to distance between them, $V(r) \propto r$. Now, the energy can be converted into equivalent mass and vice-versa as per Einstein's famous mass energy equivalence formula, $E = mc^2$. The converted mass has to be in the form of pairs of particle and its antiparticle so that the fundamental charge conservation law is conserved. A pictorial representation of one such process is given in fig. 1. At some point while stretching, it is more energetically favorable for a new quark-antiquark pair to spontaneously appear from a given energy, than to allow the bag to extend further. So at this point, the bag does not stretch further instead new quark-antiquark pairs appear from all the energy given so far for stretching the bag. These pairs destroy the original bag and again clump together in new bags, thus forming mesons and baryons. The whole process is shown pictorially in fig. 2. A collision of particles in colliders can be imagined as this process. When bags (hadrons) collide, they undergo deformation; at the energy where the pair-production is more favorable than further deformation, original bags are broken and different new bags (mesons and baryons) are created, which move away giving rise to jets. So in this way, The Bag Model correctly dummies hadrons and what might be happening at the time of collisions of hadrons so that we don't get to see individual quarks.

4 Conclusion

We explained the quark confinement and the physics responsible for it, and discussed how and why this phenomenon is drastically different from our real world experience using an example. Then we described The Bag Model which helps understand the confinement phenomenon easily.

References

[1] M. Gell-Mann, *phys. Lett.* **8**, 214-215 (1964) George Zweig, Unpublished

[2] A. Chodos, R. L. Jaffe, K. Johnson and C. B. Thorn, *Phys. Rev. D* **10**, 8 (1974)

Suggested Readings:

- David J. Griffiths, *Introduction to Elementary Particles*, John Wiley sons, New York, pp.1-78, 2004

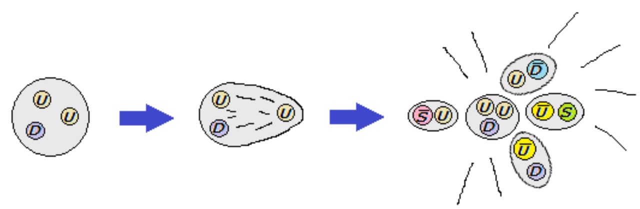


Figure 2: Bag model of quark confinement

Foucault's Pendulum

Dr. S. S. VERMA

Department of Physics
S.L.I.E.T., Longowal; Distt.-Sangrur (Punjab)-148 106

(Submitted 04-11-2014)

Abstract

Science students always feel curious to learn and also visualize the revolution of earth around its own axis in 24hrs. The introduction of the Foucault pendulum in 1851 somehow overcame all the bottlenecks and aligned itself with the Rest of the Universe and was the first dynamical proof of the rotation of earth in an easy-to-see experiment and it created a justified sensation in both the learned and everyday worlds. Today, Foucault pendulums are popular displays in science museums and universities world over. This article, briefly presents the topics like history, mechanics, construction and setting up about Foucault's pendulum with a hope that it will act as a source for its readers to understand the physics of Foucault's pendulum in a simple manner.

1. Introduction

People already knew, at the time, that the Earth turned around its own axis in 24hrs and also goes around the Sun, once every year. The Sun, in turn, goes around the centre of our galaxy, The Milky Way, once every 250 million years. These are all local motions. The cycle of night and day was already very convincing proof of earth's rotation about its own axis but the French physicist Léon Foucault wanted to prove it without resorting to any celestial elements that had already been done 218 years earlier by Galileo. The introduction of the Foucault pendulum in 1851 somehow ignores all these local motions and somehow aligns itself with the Rest of the Universe and was the first dynamical proof of the rotation of earth in an easy-to-see experiment and it created a justified sensation in both the learned and everyday worlds. Today, Foucault pendulums are popular displays in science museums and universities world over and author has the opportunity to see the Foucault pendulum installed at Physical Research Laboratory (PRL) Ahmadabad (Gujarat). Its action is a result of the Coriolis effect. It is a tall pendulum free to oscillate in any vertical plane and ideally should include some sort of motor so that it can run continuously rather than have its motion damped by air resistance.

Main installation

Pendulum bob

2. History

Mr. Foucault—a surgeon by training, physicist, inventor and journalist by trade—was not looking for direct proof that the planet rotates. The Foucault pendulum was invented by accident. In 1848 Leon Foucault was setting up a long, skinny metal rod in his lathe. He "twanged" it, and the end of the piece of metal proceeded to go up-and-down. If you treat the chuck of the lathe like a clock, the end vibrated from 12 o'clock down to 6 o'clock, and back to 12 o'clock, and so on. He slowly rotated the chuck by 90 degrees. But the end of the metal rod steadfastly vibrated back-and-forth between 12 and 6 o'clock! This set Leon Foucault thinking. He set up a small pendulum in his drill press. He set the pendulum oscillating, and then started the drill press. Once again, the pendulum kept swinging in its original plane, and ignored the fact that its mounting point was rotating. He then constructed a 2 metre-long pendulum with a 5 kilogram ball in his workshop in his cellar. Before the amplitude of the swing died away totally, he saw that the weight on the end of the pendulum appeared to rotate clockwise .

Now that he was convinced of the principle, he built a second pendulum with an 11-metre wire in the Paris Observatory, and it too rotated clockwise. He was asked to construct something "big" for the 1850 Paris Exhibition, and he constructed a 67-metre tall Foucault pendulum in the PanthŽon - a Parisian church also known as the church of Saint Genevi ve. He went to a great deal of trouble to make sure that the wire was perfectly symmetrical in its metallurgy. He used a 28 kilogram cannon ball. A stylus was placed under the ball, and sand was scattered under the potential path of the ball, so that the stylus would cut a trace in the sand. The ball was pulled to one side, and held in place with a string. With much ceremony, the string was set alight, and the ball began to describe a beautiful, straight (non-elliptical) path in the sand. Within a few minutes, the pendulum had begun to swing a little clockwise - and the previous, narrow straight-line in the sand had widened to look like a twin-bladed propeller. The experiment was a success. The Earth rotated "under" his pendulum. So it was possible, way back in 1850, to set up an experiment inside a room which had no view of the outside world, and prove that the Earth rotated. The next year, Foucault repeated his Pendulum experiment with a massive, spinning weight. He showed that this weight, just like his Pendulum, ignored local effects and lined itself up with the distant stars.

During Foucault's life it was already proven that the Earth rotated; thanks to experiments which showed that weights dropped from tall towers fell slightly to one side rather than straight down. This is the Coriolis Effect in action. Essentially, the weight does fall in a straight line towards the centre of the Earth, as you would expect, but the Earth rotates slightly during its fall. It is the tower and the ground which move sideways, not the weight. Incidentally, pilots on long flights have to correct for this effect; if they took off, pointed their plane at a distant destination and maintained a straight course, their arrival airport would no longer be there when they arrived.

The first public exhibition of a Foucault pendulum took place in February 1851 in the Meridian of the Paris Observatory. A few weeks later Foucault made his most famous pendulum when he suspended a 28 kg brass-coated lead bob with a 67 meter long wire from the dome of the Panthéon, Paris. The plane of the pendulum's swing rotated clockwise 11° per hour, making a full circle in 32.7 hours. The original bob used in 1851 at the Panthéon was moved in 1855 to the Conservatoire des Arts et Métiers in Paris. A second temporary installation was made for the 50th anniversary in 1902. During museum reconstruction in the 1990s the original pendulum was temporarily displayed at the Panthéon (1995), but was later returned to the Musée des Arts et Métiers before it reopened in 2000. On April 6, 2010 the cable suspending the bob in the Musée des Arts et Métiers snapped, causing irreparable damage to the pendulum and to the marble flooring of the museum. An exact copy of the original pendulum has been swinging permanently since 1995 under the dome of the Panthéon, Paris.

3. Explanation of mechanics

Newton's law of motion states that when a body is set in motion, it will move in a straight line from its origin, as long as it is not influenced by outside forces. This was the concept upon which Foucault based his proof that the earth rotates. If you start a Foucault pendulum swinging in one direction, after a few hours you will notice that it is swinging in a quite different direction. The earth, on the other hand, will rotate once every 24 hours underneath the pendulum. Thus if you stood watching the pendulum, after a quarter of an hour or so, you would be likely to notice that the line of the pendulum's swing has changed to a different direction. This would be especially clear if one marked the position of the line of swing in the morning and had the pendulum knocking down pegs arranged in a ring at the center. However, if you are standing on the floor of a building housing a pendulum (which is connected to the earth), you will naturally think that the floor is stable and the pendulum is moving. This is because we naturally assume that the base on which we stand is stable unless our eyes or sense of balance tells us otherwise. If our base moves slowly or accelerates smoothly, we are easily fooled into thinking that another object we see is moving. Thus, after thinking for a while about the total situation you might be willing to agree that what you are seeing is a real demonstration that the earth is rotating under the pendulum and that the line of swing of the pendulum just appears to rotate.

The Earth's rotation causes the trajectory of the pendulum to change over time, knocking down pins at different positions (or oscillating in different marked directions in a circle) as time elapses and the Earth rotates. The experimental apparatus consists of a tall pendulum free to swing in any vertical plane. The actual plane of swing appears to rotate relative to the Earth. The wire needs to be as long as possible—lengths of 12–30 m (40–100 ft) are common. At either the North Pole or South Pole, the plane of oscillation of a pendulum remains fixed relative to the distant masses of the universe while Earth rotates underneath it, taking one sidereal day to complete a rotation. A pendulum day is the time needed for the plane of a freely suspended Foucault pendulum to complete an apparent rotation about the local vertical. This is one sidereal day divided by the sine of the latitude. So, relative to Earth, the plane of oscillation of a pendulum at the North Pole undergoes a full clockwise rotation during one day; a pendulum at the South Pole rotates counterclockwise. When a Foucault pendulum is suspended at the equator, the plane of oscillation remains fixed relative to Earth. At other latitudes, the plane of oscillation precesses relative to Earth, but slower than at the pole; the angular speed, ω(measured in clockwise degrees per sidereal day), is proportional to the sine of the latitude, φ ($\omega = 360 \sin\varphi$ °/day), where latitudes north and south of the equator are defined as positive and negative, respectively.

4. Construction and setting up

The Foucault pendulum (support + wire + iron ball) is attached to this building. Any pendulum consists of a cable or wire or string and a bob. For a pendulum to easily demonstrate the Foucault effect, it should have as long a cable as possible (this one is 52 feet) and a heavy symmetrical bob (this one is hollow brass, weighing about 240 pounds). Like all pendulums this one loses a bit of energy with each swing due to friction from air currents and vibrations in the cable and other factors. Thus, left to itself the pendulum would swing in shorter and shorter arcs until after a few hours it will decrease almost to zero. To keep the Foucault pendulum going, one must replace the energy lost with each swing. This can be done by giving the pendulum a little "kick" with each swing. To do this, two iron collars are attached to the cable near the top. There is a doughnut-shaped electromagnet built into the ceiling, and the iron collar swings back and forth inside the hole of the doughnut. When the pendulum cable reaches a particular point in its swing, it is detected by an electronic device and the magnet is turned on at just the right time to give the collar (and thus the cable and the bob) a little "kick" in the exact direction of its natural swing. This restores the energy lost during the swing and keeps the pendulum from stopping. It has no effect on the direction of the swing, and thus does not interfere with the demonstration that the earth is rotating.

At either the North Pole or South Pole, the plane of oscillation of a pendulum remains pointing in the same direction while the Earth rotates underneath it, taking one sidereal day to complete a rotation. When a Foucault pendulum is suspended somewhere on the equator, then the plane of oscillation of the Foucault pendulum is at all times co-rotating with the rotation of the Earth. What happens at other latitudes is a combination of these two effects. At the equator the equilibrium position of the pendulum is in a direction that is perpendicular to the Earth's axis of rotation. Because of that, the plane of oscillation is co-rotating with the Earth. Away from the equator the co-rotating with the Earth is diminished. Between the poles and the equator the plane of oscillation is rotating both with respect to the stars and with respect to the Earth. The direction of the plane of oscillation of a pendulum with respect to the Earth rotates with an angular speed proportional to the sine of its latitude; thus one at 45° rotates once every 1.4 days and one at 30° every 2 days.

A Foucault pendulum is tricky to set up because imprecise construction can cause additional veering which masks the terrestrial effect. Air resistance damps the oscillation, so Foucault pendulums in museums usually incorporate an electromagnetic or other drive to keep the bob swinging. Its oscillations continue with reassuring regularity, its movement maintained by an electromagnetic device in the base of the installation. Once a pendulum has been set in motion, it does not change the direction in which it swings. The pendulum is only attached to the Earth at a single point, which allows it to maintain its direction of oscillation. Another phenomenon, then, must be responsible for the progressive fall of the pins. Indeed, it is not the pendulum that turns, but the Earth, turning beneath the pendulum. The rotation of the instrument is only apparent; it is us, the building, the table and little pins, all firmly attached to the Earth, that are turning around the pendulum.

5. Precautions

A Foucault pendulum requires care to set up because imprecise construction can cause additional veering which masks the terrestrial effect. The initial launch of the pendulum is critical; the traditional way to do this is to use a flame to burn through a thread which temporarily holds the bob in its starting position, thus avoiding unwanted sideways motion. Air resistance damps the oscillation, so some Foucault pendulums in museums incorporate an electromagnetic or other drive to keep the bob swinging; others are restarted regularly, sometimes with a launching ceremony as an added attraction.

6. Conclusion

We all know that the Earth rotates, even if we rarely contemplate the fact, but watching a Foucault pendulum is a humbling reminder that we're all on the surface of a planet spinning in space. That intuitive impact is exactly why this experiment remains so famous. Mechanically, it’s one of the simplest experiments possible: a heavy weight attached to a very long string or cable that is free to swing in any vertical plane. This pendulum is set in motion very carefully to avoid introducing any sideways motion, usually by tying it back with a thread of cotton which is then burned with a candle. Oddly, it will then appear to change its direction of swing over time without any outside input. Of course, it is actually the Earth which is rotating, while the pendulum continues to swing in the same plane relative to the rest of the Universe. The Foucault pendulum is currently working. The main problems involved in making a Foucault Pendulum are starting the ball in a swing that passes through the true centre point of the swing, keeping the ball in that "true" swing (and not going into an elliptical swing), and pumping energy into the swing so that it does not die down. Three subsystems were set up to achieve these goals, and currently, they are all operational.

References:

[1] S. S. Marsden, Scientific American, 115(1958).
[2] J. B. Hart, R. E. Miller and R. L. Mills, Am. J. Phys. **55**, 67(1987).
[3] http://www.calacademy.org/products/pendulum/
[4]http://www.animations.physics.unsw.edu.au//jw/foucault_pendulum.html
[5]http://www.si.edu/Encyclopedia_SI/nmah/pendulum.htm
[6]http://sciencenetlinks.com/lessons/foucaults-pendulum/

Solar flare and its interaction with the Earth atmosphere: An Introduction

Pramod Kumar[1], Yogesh C. Bhatt [2], Rajmal Jain[2*], Yagvendra S. Shishodia[1]

[1, 2]Department of Physics, Jagan Nath University, Jaipur, 303901, India
[2]Malaviya National Institute of Technology Jaipur, 302017, India

[2*]Kadi Sarva Vishwavidyalaya, Gandhinagar, 382016, India
and Physical Research Laboratory, Ahmedabad, 380009, India

(*Submitted* 08-09-2014)

Abstract

Solar flare is a diverse dynamic activity of solar atmosphere. It is associated with high magnetic energy release (10^{27}J), particle acceleration, radiation burst and release of plasma particles into the space. In this paper we studied the different aspects of solar flare such as magnetic reconnection, radiation burst, magneto hydrodynamic, and interaction with Earth atmosphere from the fundamental point of view.

1. Introduction

Sun is a main sequence G2 spectral class star of the universe and full of mysterious activity like Sun Spots, Solar Cycles, Solar Flare, CME, Solar Prominences etc. Sun is a gaseous ball of plasma and continuously emitting radiations. The total solar radiation of the entire Sun is 3.83 X 10^{23} kW [1]. Sun is a magnetic active star [2] and mostly activity is due to the variation in magnetic field. Solar magnetic field is generated in the convection zone by a dynamo process which is a combination of solar rotation and convection [3]. The magnetic field of various activity of Sun are depicted in table 1[1]. According to solar standard model [4], Sun structure is considered as an internal as well as external structure (solar atmosphere). Solar atmosphere comprises photosphere, chromospheres and corona. The basic temperature and density parameters of Sun structure are depicted in table 2 [1].

Solar flare is an solar atmospheric activity and first discovered or observed by R. C. Carrington and R. Hodgman on 1st September 1859 independently in optical light [5][6]. Solar Flare is associated with solar spots (high magnetic field, ~3000 Gauss), solar cycle (variation in number of sun spots in 11 year) and sometime by coronal mass ejections (plasma particles propagation along with magnetic field lines). Solar Flare is an intense variation of energy into solar corona. Though, Sun is continuously emitting energy in the form of radiation in all wavelengths from radio to gamma rays (please see Appendix A for wavelengths, frequency and energy of electromagnetic radiation spectrum) but during flare an intense increase arise in the radiation level of the Sun [7]. Also huge amount of plasma particles accelerate in the space along with magnetic field lines say CME (sometimes), solar wind etc.

It is widely believed that flares are caused by the release of magnetic energy up to some 10^{27} J in the solar atmosphere within a relatively short time between 100 and 1000 s. Release of magnetic energy causes particle acceleration (electrons, protons, and heavier particles) and turbulence in solar atmosphere. This energy accelerates particles to tens and hundreds of keV. Electrons accelerated to 10-100 keV can contain a significant fraction (10-50%) of this energy. The accelerated particle causes intense radiations in all wavelengths. Inside the solar flare, gas temperature can reach up to 10 or 20 million degrees Kelvin. It can be as high as 100 million degrees Kelvin.

It is observed that there are typically three stages in a solar flare [8]. First is the precursor stage, where the release of magnetic energy is triggered. Soft X-ray emission is detected in this stage. In the second or *i*mpulsive stage, protons and electrons are accelerated to energies exceeding 1 MeV. During the impulsive stage, radio waves, hard X-rays, and gamma rays are emitted. The gradual build up and decay of soft X-rays can be detected in the third, decay stage. The duration of these stages can be as short as a few seconds or as long as an hour**.**

Solar flares are mainly categorized in single loop and two ribbon flares. Single loop flares are single magnetic loop or flux tube brightens in X-rays and remains apparently unchanged in shape and position throughout the event. Two ribbon flares are much larger than a compact flare and takes place near a Solar prominence, a loop of plasma confined between two magnetic field lines [9]. The Hα and Soft X-ray classification scheme is given in appendix A. A B C M and X class classification depend upon peak flux. The X-class flare denotes the most intense flares, while the number provides more information about its strength. An X2 is twice as intense as an X1, an X3 is three times as intense, etc.

In the present time, solar flares are observed in all wavelengths from radio to gamma rays emissions excess in to 10 MeV. Space based observatories and Ground based observatories are in use for observing solar flare. YOHKOH, GOES, RHESSI, SDO, SOHO, WIND was/are the major solar space observatories [10]. An example of solar flare observed by SOHO spacecraft is shown in figure 1. A flare observed by the GOES satellite based on the X-ray radiation level of Sun is shown in figure 2.

In this paper we studied introductory aspects of solar flare, magnetic reconnection, magneto hydrodynamic, radiative process and its interactions with the Earth atmosphere.

Table 1: Magnetic field of different activities on Sun

Solar Activity	**Magnetic Field (Gauss)**
Polar Field	1-2
Sun Spots	3000
Promineces	10-100
Chromospheric plages	200
Bright chromospheric network	25
Ephemeral (unipolar) active regions	20

Table 2: Temperature and density parameters of Sun Structures

Solar Structure	**Temperature (K)**	**Density (Kg/m3)**
Interior (centre of Sun)	15 000 000 K	16000
Surface (photosphere	6050 K	10^{-6} kg/m3
Sunspot umbra (typical)	4240 K	
Sun spot Penumbra (typical)	5680 K	

Chromospheres	4300 to 50 000 K	10^{-9} kg/m3
Corona	800 000 to 3 000 000 K	10^{-13} kg/m3

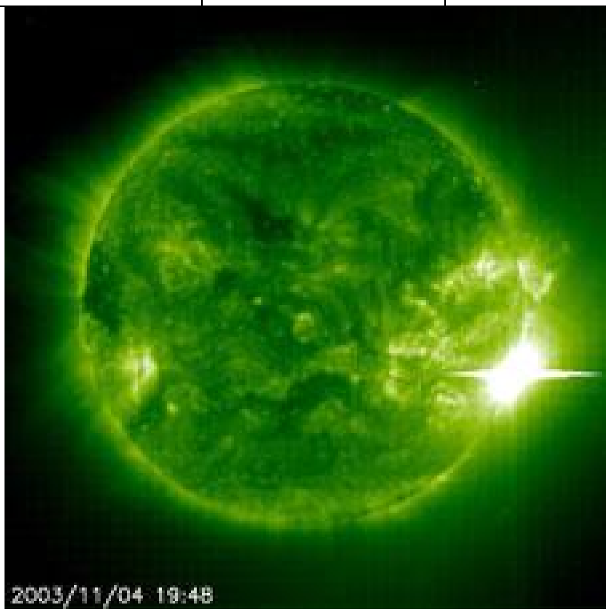


Figure 1: The Sun unleashed a powerful flare on 4 November 2003. The Extreme ultraviolet Imager in the 195A emission line aboard the SOHO spacecraft captured the event. **Credit:** SOHO, ESA & NASA

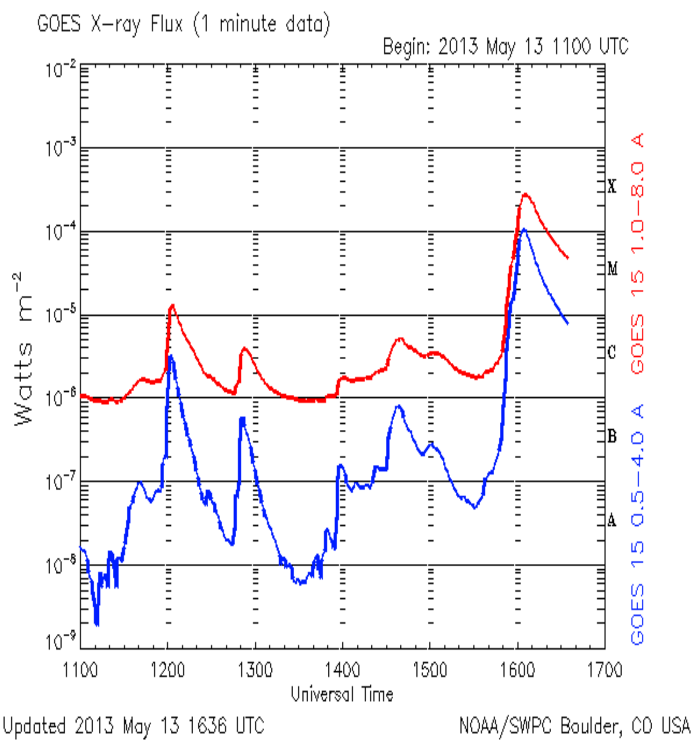


Figure 2: X-ray flux from the Sun of 13.05.2013 (GOES-15), Image courtesy: NASA

2. Magnetic Reconnection

Magnetic reconnection is a natural process of plasma state of the matter. Magnetic reconnection is a process where magnetic field lines are broken and rejoined in highly conducting plasma. Solar flare is caused by magnetic reconnection. During solar flare magnetic reconnection will dissipate magnetic energy which will cause particle acceleration and release of huge amount of emissions. The process of magnetic reconnection was first postulated by Giovanelli (1946) [11].

The term magnetic reconnection was first proposed by Dungi in 1953[12]. Dungy suggested the formation of current sheet within the perspective of magneto hydrodynamics. In the current sheet the magnetic field will reconnected and diffusion occurs. The differential rotation of Sun, solar dynamo, and convection ensure that the magnetic field of the Sun is typically highly stressed and prone to reconnection events [13]. Here we presented cartoons for solar flare formation taking magnetic reconnection as a process. It is shown in figure 3 a, 3b, 3c and 3d respectively. The figures are not in scale.

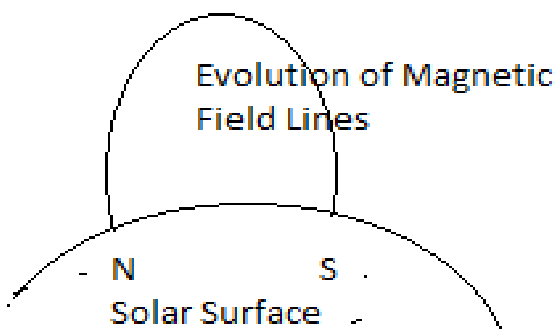


Figure 3a: Evolution of magnetic field lines

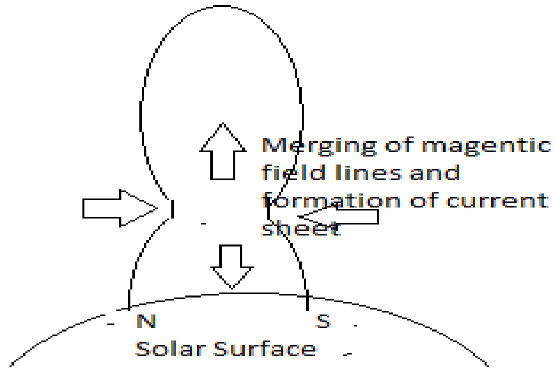


Figure 3b: Twisting of magnetic field lines and current sheet formation

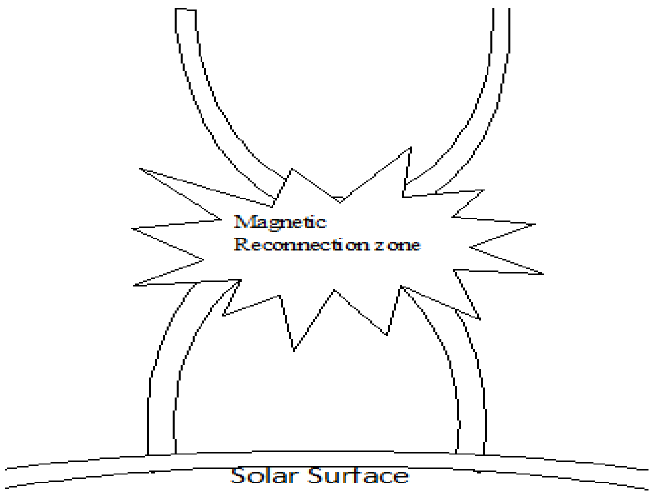


Figure 3c: Release of magnetic Energy

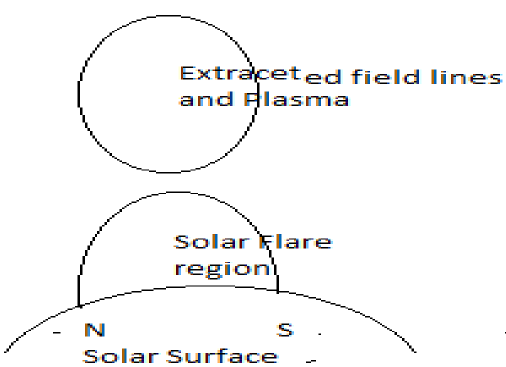


Figure 3d: Solar flare formation and extraction of field lines

After the postulation of magnetic reconnection several models have been proposed. The first model to describe magnetic reconnection was developed by Parker (1957) and Sweet (1958) [14] [15] in terms of enhanced magnetic diffusion in a layer with anti parallel field lines on both sides. This model found slow at time scale in solar flare conditions. Later Petschek [16] develop fast reconnection model considering small current sheet. A schematic of Sweet Parker and Petschek models of Reconnection is shown in figure 4

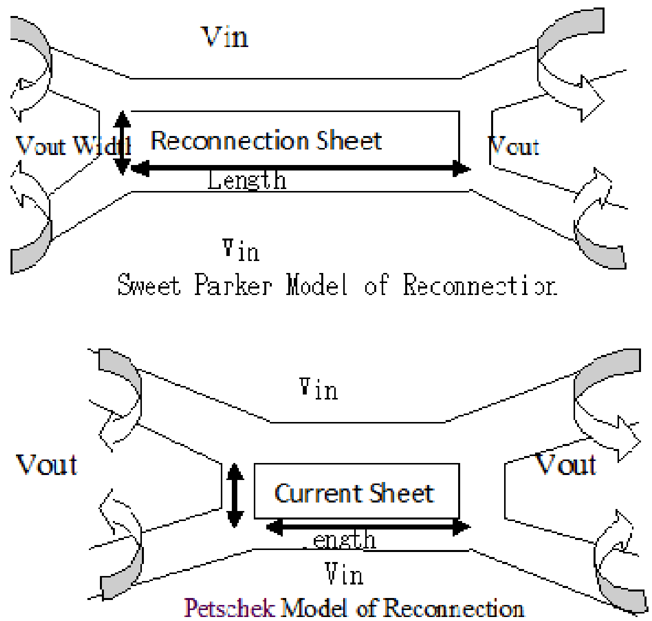


Figure 4: Sweet-Parker and Petschek Model of magnetic reconnection (figures are not in scale)

Under steady state conditions Sweet parker calculated the rate of flow of plasma from the current sheet. If E_y is the uniform out of plane electric field and B_{in} is the magnetic field of reconnection region, than;

$$E_y \sim V_{in} B_{in}$$

According to Appere's law, the current density is given by;

$$J_y \sim \frac{B_{in}}{\mu_o \delta}$$

Where δ is the half thickness of Sweet parker model thickness. A relation resistivity η, in flow velocity V_{in} and current sheet half thickness δ will be equal to;

$$V_{in} \sim \frac{\eta}{\mu_o \delta}$$

The above relationship represents the condition that the resistive electric field, ηJy, within the current sheet matches the ideal electric field outside the current sheet. The thickness of a steady state reconnection layer is set by how quickly magnetic field lines can diffuse.

Conservation of mass will give a relationship under incompressible conditions;

$$V_{in} L \sim V_{out} \delta$$

Where L is the half thickness of the reconnection layer and Vout is the outflow velocity. According to the conservation law of energy;

$$V_{in} L \left(\frac{B^2{}_{in}}{2\mu_0} \right) \sim V_{out} \delta \left(\frac{\rho V^2{}_{out}}{2} \right)$$

Solving this equation we will get following relationship

$$V_{out} \sim V_A \equiv \frac{B_{in}}{\sqrt{\mu_0 \rho}}$$

Thus out flow velocity in a Sweet Parker current sheet scales with the upstream Alfen speed. The dimension less reconnection rate ca be given by

$$\frac{V_{in}}{V_A} \sim \frac{1}{\sqrt{S}}$$

Where S is the Lundquist number and is equal to;

$$S = \frac{\mu_0 L V_A}{\eta}$$

Sweet Parker model is not able to explain the extremely short time scales of seconds to minutes observed for energy release during solar flare. Sweet Parker is found correct for the resistive MHD framework for moderate Lundquist number ($S \sim 10^2 - 10^4$).

The second most important model was proposed by Petschek (1964)[16] as shown in figure 4. Petschek modify Sweet parker model considering small current sheet. Perschek considered that diffusion region must be small than the global length scale. He proposed a mechanism for the diffusion in a smaller length scale considering slow mode shocks. The maximum reconnection rate is given by

$$\frac{V_{in}}{V_{out}} \approx \frac{\pi}{8 \ln S}$$

In the beginning, the Petschek model got good acceptance but it is found that under resistive MHD conditions only Sweet Parker type current sheet will form as studied by Biskamap in 1986 [17]. Experimental and observational results ruled out the formation of Petschek current sheet.

Apart from these two models several other models are also given by researchers such as Turbulent Reconnection, Two Fluid Reconnection, Asymmetric Reconnection etc.[18].

Multiple wavelength observations by modern satellites strongly suggest that magnetic reconnection is the principal process for the energy conversion in solar flares. Typical cusp-like feature of post-flare loops observed in soft X-rays is a strong piece of evidence in favour of the reconnection model [19].

The solar flare standard model [20] is known as CSHKP model. This model was developed by pioneer researchers Carmichael (1964), Sturrock (1966), Hirayama (1974), and Kopp-Pneuman (1976). Kopp and Pneuman (1976) considered that after reconnection of open field line, the solar wind along open field line collides to form shock inside the reconnected closed field, which heat the coronal plasma to are temperature. However, Cargill and Priest (1982) correctly pointed out that we should consider the role of slow mode shock associated with Petschek type reconnection. Forbes and Priest (1984) noted the formation of fast shock (termination shock) due to reconnection jet above the reconnected loop, and Forbes and Malherbe (1986) pointed out that the slow shock is dissociated to isothermal slow shock and conduction front in solar are condition.

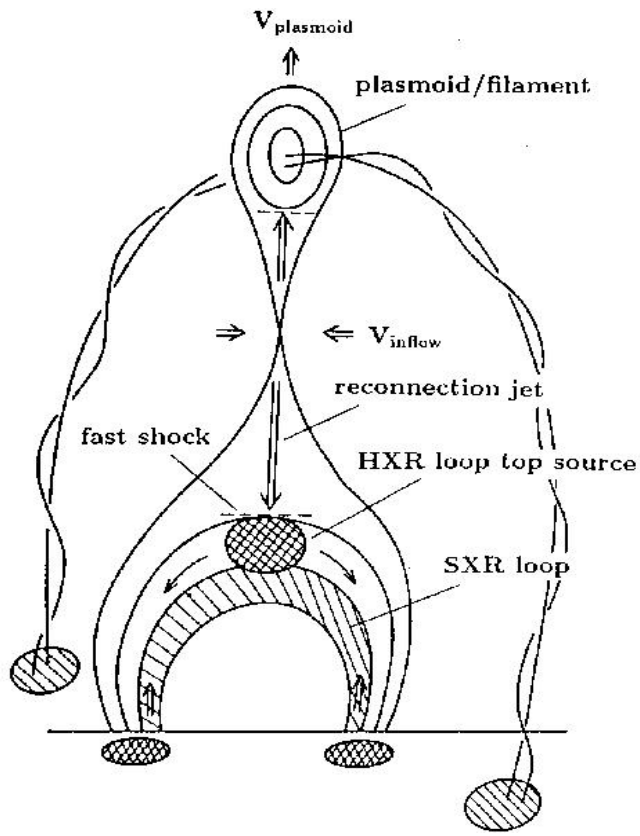


Figure 5: A standard model of solar flare as developed by Shibata

Some of the solar flares parameters observed and derived from the theoretical modeling are summarized in table 3[7] [13].

Table 3: Physical parameters of the Solar Flare

Flare Properties	**Values**
Flare Energy	10^{27} J
Flare Temperature	T0^{7} K
Flare Volume	$(10^4\ \text{km})^3$
Fare Loop Height	10^4 to 10^5 Km
Plasma Density	$10^{10}\ \text{cm}^{-3}$
Duration of flare (time scale)	$10^3 - 10^4$ s
Electron energies	>10 MeV
Proton energies	>100 MeV
Number of energetic electrons	10^{36} per second

3. Magneto hydrodynamics of the Flare

Magneto hydrodynamic is a dynamics of plasma material in the influence of magnetic field. Magneto hydrodynamics is similar to fluid dynamics. Magneto hydrodynamic equation, relates different parameters of plasma, are used for studying dynamics of plasma. The solutions of MHD equations provide detail information about the plasma flow. Analytical and Numerical methods are used for solving magneto hydrodynamic equations.

MHD equations are non-linear partial differential equations. The basic Magneto hydrodynamic equations [21] are as follows:
Mass conservation equations;

$$\frac{\partial\rho}{\partial t}+\nabla.(\rho V)=0$$

Momentum conservation equation;

$$\rho\frac{\partial V}{\partial t}+\rho(V.\nabla)V=-\nabla p+\tfrac{1}{\mu}(\nabla\times B)\times B+\rho g$$

Faraday's equations

$$\frac{\partial B}{\partial t}+(\nabla\times E)=0$$

Gauss law;

$$\nabla.B=0$$

Ampere's law;

$$\nabla\times B=\mu j$$

Equation of state;

$$p = \frac{\kappa_B}{m}\rho T$$

Here ρ is mass density, V is flow velocity, B is the magnetic field, E is the electric field, p is gas pressure, $\gamma = 5/3$ is the adiabatic index, g is gravitational acceleration, T is temperature, m is mean particle mass and k_B is the Boltzmann's constant. MHD equations relates plasma unknown parameters with one another.

Induction equation is derived by using Ampere's law and ohm's law,

$$\frac{\partial B}{\partial t} = \nabla\times(v\times B) + \eta\nabla^2 B$$

The first term of the right side describes advection and second term diffusion.

A dimension less number, Magnetic Reynold number is defined as a ratio of advection and diffusion terms.

$$R_m = \frac{\nabla\times(v\times B)}{\eta\nabla^2 B}$$

If $R_m \gg 1$; advection term will dominate (condition is known as frozen in flux).

If $R_m \ll 1$; Diffusion term dominate and reconnection occurs.

Plasma beta parameter is defined as a gas pressure (p) to magnetic pressure ($B^2/2\mu$). It is equal to;

$$\beta = \frac{2\mu p}{B^2}$$

If $\beta \gg 1$, the gas pressure will dominate (Photosphere of Sun)

If $\beta \ll 1$, the magnetic pressure will dominate (Solar corona)

Forbes & Priest 1983; Magara et al. 1996; Ugai 1996; Forbes & Malherbe 1991 carried out magneto hydrodynamic simulation on magnetic reconnection models in two dimensions. This simulation work was mainly concentrated to simulate basic geometry of solar flare such as the reconnection site structures with an X-type magnetic neutral point associated with extending slow-mode MHD shocks, bidirectional reconnection jets, fast-mode MHD shock formed at the top of reconnected loops, and upward ejecting plasmoids. These simulations do not include thermal processes such as heat conduction. Forbes and Melherbe ((1991) simulated magnetic reconnection by considering radiative cooling effect.

Yokoyama and Shibata (1997) carried out first time the self-consistent MHD simulation for the reconnection. Yokoyama and Shibata (1998, 2001) [22] [23] performed 2D MHD simulation of reconnection with heat conduction and chromospheres evaporation. This simulation is considered as a most advanced model of eruptive areas.

An illustration of the reconnection model of a solar flare based on the simulation results is shown in figure 6. This model was develop by the simulation study carried out by the Petschek 1964; Tsuneta 1996; Shibata 1996. Thick solid lines show magnetic field. Magnetic energy is released at slow-mode MHD shocks emanating from the neutral X-point, which is formed as a result of the magnetic reconnection. The ejected reconnection jet collides with the reconnected loops and forms a fast-mode MHD shock. The released heat at the reconnection site conducts along the field lines down to the chromospheres. Because of the heat input into the dense chromospheres plasma, the plasma there evaporates and flows back toward the corona.

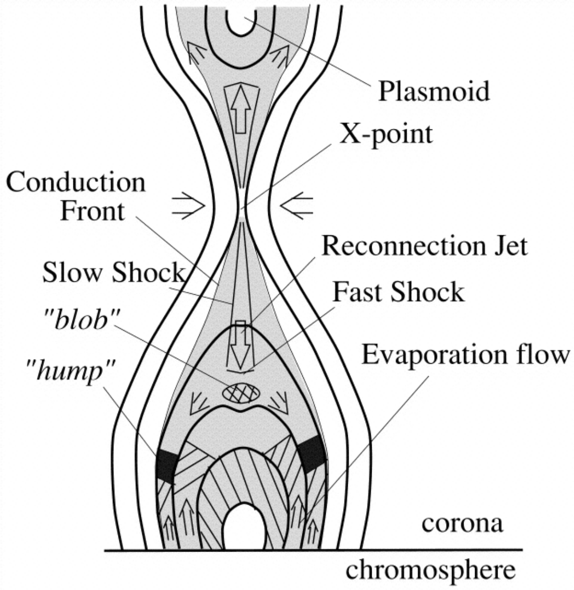


Figure 6: Schematic picture of the model; e.g., Petschek 1964; Tsuneta 1996; Shibata 1996

4. Flare Radiations

Most of information about the astronomical bodies, phenomenons and processes are obtained by studying electromagnetic radiations which comes from the object. Radio, microwave infrared, visible, ultraviolet, X rays, gamma rays are known as electromagnetic radiations. The electromagnetic (EM) spectrum is the range of all types of EM radiation. The wavelength, frequency and energy of electromagnetic radiation are given in the Appendix. The basic properties associated with radiation are luminosity, flux, intensity, emissivity, radiation energy density, Einstein coefficients, mean free path etc. **[24]**

Electromagnetic radiation is produced whenever a charged particle is accelerated. The greater the acceleration the higher the energy of the emitted photon will produce. The basic mechanisms for the production electromagnetic radiations are as follows:

- Bremsstrahlung (free-free emission)
- Compton Scattering
- Synchrotron emission
- Absorption processes
    - Self-absorption
    - Pair production
- Ionisation Losses
- Cherenkov radiation
- Nuclear interactions

There are two types of electromagnetic radiation: thermal radiation - which depends on the temperature of the emitting source - and non-thermal - which does not depend on the source temperature. The thermal radiation processes are black body radiation and thermal bremstrahlung. The Non thermal emissions are non thermal bremstrahlung, inverse Compton scattering and synchrotron radiation

X-ray and radio emission mechanism can also be classified as coherent and incoherent emissions. Coherent emission mechanisms are plasma emission and the electron cyclotron maser mechanisms. Incoherent emission mechanisms include bremsstrahlung and gyro-emission.

In solar flare an intense radiation occurs in all wavelength of electromagnetic spectrum at different Locations of flare regions. Nearly half of energy released during flares is used to accelerate electrons and protons up to velocity nearly speed of light. Solar flares accelerate particles to nearly the velocity of light, hurling them out into the solar system and down into the Sun. Both Thermal and Non thermal radiations occurs in solar flares [26].

X-ray emissions from solar flare provide a detail study of physical properties of Solar flare. X-ray emission can be classified as hard X-ray (10-100KeV) emissions and soft X-ray (~0.1-10 keV) emissions. SXR often refers to the thermal part of the photon bremsstrahlung spectrum, which can go up to 20 keV in powerful flares. Hard X-ray referred as non thermal power law like spectrum. The hard X-rays are widely believed to be non thermal bremsstrahlung emission produced by high-energy electrons precipitating into the chromospheres. High-speed electrons that are thrown down into the Sun emit hard X-rays when entering the lower solar atmosphere [26]. An illustration of X-ray photon production in solar flare is shown in figure 7.

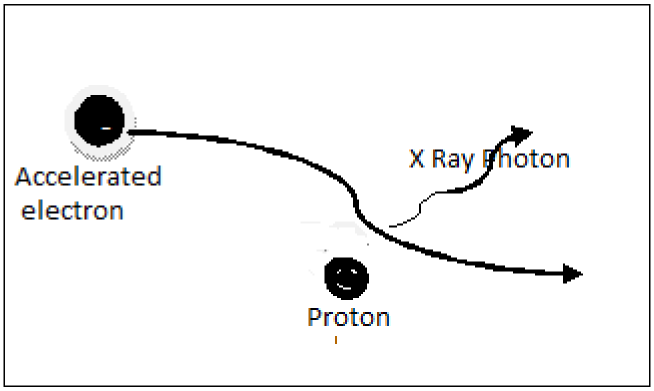


Figure 7: Production of X ray in Solar Flare

Hard X-ray emission and the evolution of high-temperature thermal flare plasma are often referred to as the Neupert effect. Jain, Rajmal, et Al. [27][28][29] studied X ray emissions, observed by SOXS mission of India, from solar flare meticulously and derived many conclusions about Solar flare. According to Jain et. al, [27][28][29] The evolution of the break energy point that separates the thermal and non-thermal processes reveals increase with increasing flare plasma temperature. Jain et al., [27][28][29] also concluded that the X-ray energy spectrum from a typical large solar flare is dominated by soft X-ray line and thermal (free-free) bremsstrahlung emission at $\varepsilon \approx$ 1–20 keV, and collisional bremsstrahlung of non-thermal electrons at $\varepsilon \approx$ 20–1000 keV .

Gamma and neutrons emissions from solar flare provide information about flaring process and conditions within flare loop [30]. Gamma ray emissions are generally associated with most powerful solar flares. Energetic flare protons create nuclear reactions when they are tossed down into the chromosphere or photosphere, emitting gamma rays in the process Solar flare are generally associated with such as the nuclear lines from excited nuclei, as well as the delayed neutron capture line, and the electron-positron annihilation line and continuum.

For gamma rays, protons and heavier ions accelerated in the flare. These high energy particles interact with the nuclei of the different elements in the ambient solar atmosphere to produce a far more complicated emission spectrum than the relatively smooth continuum bremsstrahlung spectrum. Many individual gamma-ray lines from a wide variety of different elements in the solar atmosphere have been detected. They result from the decay of such relatively abundant elements as carbon, nitrogen, oxygen, etc. that are excited to high energy states in the various nuclear interactions. The relative intensities of the various lines provide information about the composition of both the accelerated particles and the target nuclei [31].

5. Earth Atmosphere and Flare Interaction

Earth atmosphere is the protective layers of gases which surrounds the Earth. It receives solar radiation in the form of radiations and plasma particles (CME, solar wind etc) from the Sun. Earth atmosphere mainly constitute nitrogen (78%), oxygen (21%) and other gases. Gravity holds the atmosphere to the Earth's surface. With the height the atmosphere becomes thinner. Earth atmospheres (figure 8) are classified into five layers;

1) The troposphere is the first layer above the surface and contains half of the Earth's atmosphere. Weather occurs in this layer. 2) Many jet aircrafts fly in the stratosphere because it is very stable. Also, the ozone layer absorbs harmful rays from the Sun. Meteors or rock fragments burn up in the mesosphere. 4) The thermosphere is a layer with auroras. It is also where the space shuttle orbits. 5) The atmosphere merges into space in the extremely thin exosphere. This is the upper limit of our atmosphere.

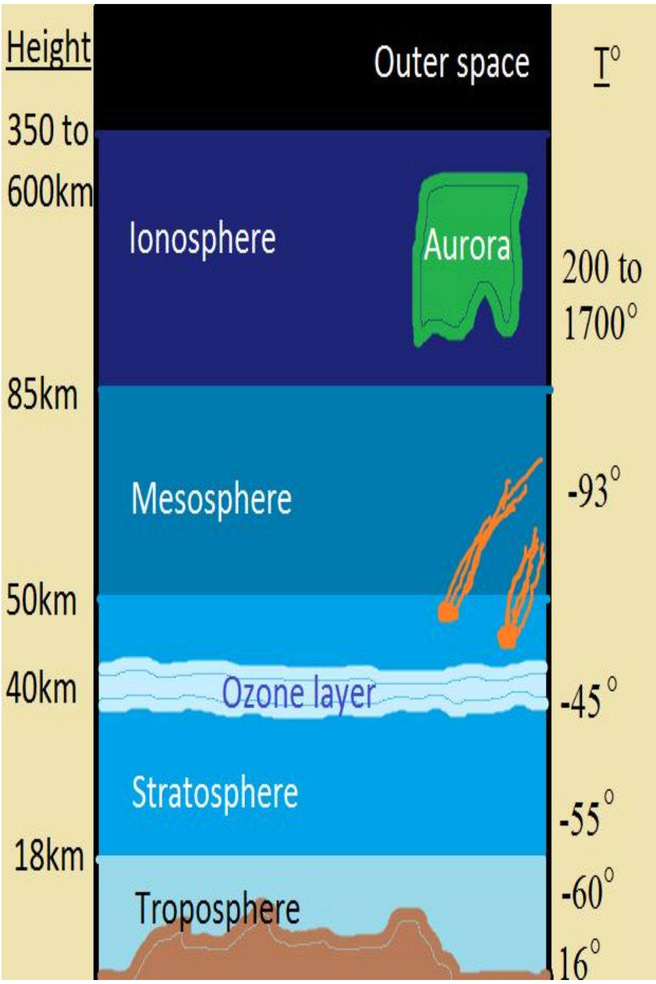


Figure 8: Earth Atmosphere with height

Flare interaction is assumed as an interaction of plasma particles and intense radiation with the Earth atmosphere. The key processes of the interaction take place are the emission, absorption and scattering. As discussed above that Solar flare produces high energy particles (protons, electrons ions etc.) and radiations. The particles are stopped by the magnetic field of earth. Radiation (mainly X ray and UV) is stopped by ionosphere layer of earth atmosphere. As the charged particles of solar winds and flares hit the Earth's magnetic field, they travel along the field lines [32].

A sudden ionosphere disturbance (SID) occurs during flare. The ionosphere starts at about 70-80 km high and continues for 640 km. It contains many ions and free electrons (plasma). The ions are created when sunlight hits atoms and tears off some electrons. Auroras occur in the ionosphere. the ionosphere is composed of three main parts, named for obscure historical reasons: the D, E, and F regions. The electron density is highest in the upper, or F region. The F region exists during both daytime and nighttime. During the day it is ionized by solar radiation, during the night by cosmic rays.

The flare's X-ray energy increases the ionization of all the layers, including the D. Electromagnetic radiation at wavelengths of 100 to 1000 Angstroms (ultraviolet) ionizes the F region, radiation at 10 to 100 Angstroms (soft X-rays) ionizes the E region, and radiation at 1 to 10 Angstroms (hard X- rays) ionizes the D region. When electromagnetic radiation from the sun strips an electron off a neutral constituent in the atmosphere, the resulting electron can spiral along a magnetic field line at the electron gyro frequency.

When a Solar Flare occurs, the VLF propagation is disturbed and it is possible to detect the SIDs by monitoring the variations or the signal level of a distant VLF receiver. Figure 9 is an artistic view of flare interaction with Earth atmosphere.

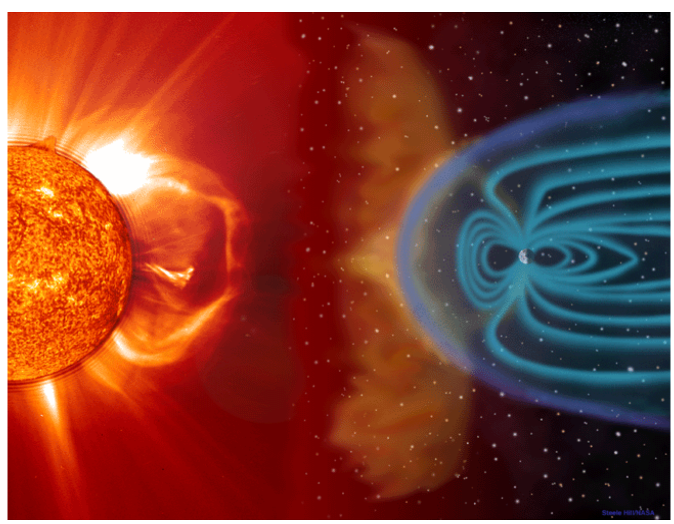


Figure 9: An artistic view of solar flare interaction with Earth atmosphere

8. Conclusions

From the beginning of the discovery [5][6] of solar flare in the solar atmosphere to till now, solar flare is studied very meticulously from the theoretical and observational point of view. Theoretical and observational studies explored solar flare very much. Theoretical study such as study of magnetic reconnection [11][12] and its models [14][15] provided very significant information about the flare which is very much correlated with the observational[10] results. Magneto hydrodynamics models and their numerical salutation [22][23] are very helpful in exploring solar flare phenomenons.

Observational study of flare by observing radiations, fundamental particles, explored solar flare very clearly. Jain, Rajmal et al. [27][28][29] studied X-ray emission from solar flare and explored solar flare from many point of view such as heavy elements presences, thermal and non thermal energy radiations etc. The study of X-ray emissions from solar flare indicates stages in solar flare. High-resolution images obtained from the hard X-ray telescope (HXT) and the soft X-ray telescope (SXT) on Yohkoh and Rhessi revealed several features of solar flare which are consistent with the reconnection model.

Now it is well known that solar flare materialize due to the release of huge magnetic energy and having magnetic reconnection [17] is the prime process behind formation. Solar flare takes place in stages, which are well correlated with observational and theoretical modeling. Solar flare has a loop like structure [17] with thermal variations at different regions in the loop. Electromagnetic radiations take place at different regions in flares. Theoretical modelings of radiations, processes, are very helpful in calculating properties, parameters and extracting information of the flare.

It is also well clear that solar flare causes increased radiation and particle dose in the space. Many phenomenons's such as Aurora, increase in ionosphere concentration level (SID), geomagnetic storms, magnetic reconnection etc. occurs in Earth outer atmosphere.

Still lots of questions are unsolved with respect to Sun and Solar flare. The detail study of Sun is not yet complete. The formation of magnetic field, hydrodynamics of coronal loops, MHD oscillations and coronal seismology, the coronal heating problem, self organized criticality from nano flares to giant flares, magnetic reconnection processes, particle acceleration processes, coronal mass ejections and coronal dimming etc. are the major unsolved problems [33]

Acknowlegements

The first author is highly thankful to his Ph.D. supervisors for their valuable guidance and encouragement throughout the writing of this paper. First author is also very thankful to all the friends for providing study materials.

References:

[1] Michael Stix, The Sun: An Introduction, (Springer, 2004).

[2] Hale, G.E., On the probable existence of a magnetic field in sunspots, The Astrophysical Journal, **28**, pps. 315-343 (1908)
.

[3] Piddington, J. H., On the origin and structure of stellar magnetic fields, *Astrophysics and Space Science* 90 (1): 217–230 (1983). Bibcode:1983Ap&SS..90..217P.doi:10.1007/BF00651562.

[4] Bahcall, J.N., Huebner, W.F., Lubow, S.H., Parker, P.D. & Ulrich, R.K., Standard solar models and the uncertainties in predicted capture rates of solar neutrinos, Reviews of Modern Physics, 54, 764, (1982).

[5] Carrington, R.C., Monthly Notices of the Royal Astronomical Society, **20**, p. 13, (1860).

[6] R. Hodgson, On a curious Appearance seen in the Sun. Mon. Not. r. Astr. Soc. 20, 15–16 (1859).

[7] Einar Tandberg-Hanssen, A. Gordon Emslie, The Physics of Solar Flares, (Cambridge University Press, 1988).

[8] B.V. Somov, Physical Processes in Solar Flares, (Springer, 1992).

[9] E. R. Priest, Solar flare magneto hydrodynamics (Gordon and Breach science publisher, 1981)

[10] Kenneth R. Lung, The Sun From Space, Springer-Verlag Berlin Heidelberg (2009)

[11] Giovanelli, R. G., Theory of Chromospheric Flares Nature, 158, 81A (1946).

[12] Dungey, J. W., Conditions for the occurrence of electrical discharges in astrophysical systems. *Phil. Mag*. 44, 725–738, (1953).

[13] Aschwanden, M. J., Physics of the Solar Corona (Chichester, UK: Springer-Praxis, 2004)

[14] Parker, E. N., Sweet's mechanism for merging magnetic fields in conducting fluids. *J. Geophys. Res*. 62, 509–520, (1957).

[15] Sweet, P. A., Mechanisms of solar flares. Ann. Rev. Astron. Astrophys. 7, 149–177 (1969)

[16] Petschek, H. E., in AAS/NASA Symposium on the Physics of Solar Flares, ed. W. N. Hess, (Washington, D.C.: NASA), 425 (1964)

[17] Biskamp, D., Phys. Fluids, 29, 1520 (1986)

[18] Murphy Nicholas Arnold Echlin, Simulation and Analysis of Magnetic Reconnection in a Laboratory Plasma Astrophysics Experiment, PHD Thesis, University of Wisconsin – Madison 2009

[19] Tsuneta, S., Structure and dynamics of magnetic reconnection in a solar flare. *Astrophys.J*. 456, 840–849, (1996).

[20] Shibata, K., Evidence of Magnetic Reconnection in Solar Flares and a Unified Modelnof Flares. Astrophysics & Space Science, 264, 129-144(1999).

[21] Eric Priest, Terry Forbes, Magnetic Reconnection: MHD theory and Applications, (Cambridge University Press, 2000)

[22] Yokoyama, Takaaki, and Kazunari Shibata. A two-dimensional magneto hydrodynamic simulation of chromospheres evaporation in a solar flare based on a magnetic reconnection model, The Astrophysical Journal Letters 494.1 (1998): L113

[23] Yokoyama, Takaaki, and Kazunari Shibata, Magneto hydrodynamic simulation of a solar flare with chromospheric evaporation effect based on the magnetic reconnection model. The Astrophysical Journal 549.2 (2001): 1160

[24] RYBICKI, G.B. & LIGHTMAN, A.P. Radiative Processes in Astrophysics, (New York: Wiley, 1979)

[25] Liu, Wei. Characteristics of solar flare hard X-ray emissions: observations and models. Diss. Stanford University, 2006.

[26] Vilmer, N., Hard X-ray emission processes in solar flares., Particle Acceleration and Trapping in Solar Flares. Springer Netherlands, 207-223, (1987).

[27] Jain, Rajmal, Malini Aggarwal, and Raghunandan Sharma. X-ray emission from solar

flares. Journal of Astrophysics and Astronomy 29.1-2, 125-145, (2008)

[28] Jain, Rajmal, et al. "Energy-Dependent Timing of Thermal Emission in Solar Flares." *Solar Physics* 270.1, 137-149, (2011)

[29] Jain, R., Rao, A. R., Deshpande, M. R., Dwivedi, B. N., Manoharan, P. K., Seetha, S., Vahia, M. N., Vats, H. O., Venkatkrishnan, P., *Bull. Astron. Soc. India*, **29**, 117, (2000a)

[30] Murphy, R., Exploring *solar flares with gamma rays and neutrons*. NAVAL RESEARCH LAB WASHINGTON DC, (2008)

[31] Gordon D. Holman, Overview of Solar Flares,http://hesperia.gsfc.nasa.gov/hessi/flares.htm, November 10, (2000).

[32] Gordon Holman, Space Weather: What impact do solar flares have on human activities? http://hesperia.gsfc.nasa.gov/sftheory/spaceweather.htm, Nov. 10 (2010)

[33] Markus J. Aschwanden, Keynote Address: Outstanding Problems in Solar Physics, J. Astrophys. Astr., 29, 3-16, (2008)

Appendix A

Approximate wavelength, frequency, and energy limits of the various regions of the electromagnetic spectrum.

	Wavelength (m)	Frequency (Hz)	Energy (J)
Radio	$> 1 \times 10^{-1}$	$< 3 \times 10^{9}$	$< 2 \times 10^{-24}$
Microwave	1×10^{-3} - 1×10^{-1}	3×10^{9} - 3×10^{11}	2×10^{-24}- 2×10^{-22}
Infrared	7×10^{-7} - 1×10^{-3}	3×10^{11} - 4×10^{14}	2×10^{-22} - 3×10^{-19}
Optical	4×10^{-7} - 7×10^{-7}	4×10^{14} - 7.5×10^{14}	3×10^{-19} - 5×10^{-19}
UV	1×10^{-8} - 4×10^{-7}	7.5×10^{14} - 3×10^{16}	5×10^{-19} - 2×10^{-17}
X-ray	1×10^{-11} - 1×10^{-8}	3×10^{16} - 3×10^{19}	2×10^{-17} - 2×10^{-14}
Gamma-ray	$< 1 \times 10^{-11}$	$> 3 \times 10^{19}$	$> 2 \times 10^{-14}$

Flare Classification Scheme

Hα classification			**Radio flux at 5000 MHz in s.f.u.**	**Soft X-ray class**	
Importance Class	**Area (Sq. Deg.)**	**Area 10^{-6} solar disk**		**Importance class**	**Peak flux in 1-8 Å w/m^2**

S	2.0	200	5	A	10^{-8} to 10^{-7}
1	2.0–5.1	200–500	30	B	10^{-7} to 10^{-6}
2	5.2–12.4	500–1200	300	C	10^{-6} to 10^{-5}
3	12.5–24.7	1200–2400	3000	M	10^{-5} to 10^{-4}
4	>24.7	>2400	3000	X	$>10^{-4}$

After Bhatnagar & Livingston 2005; Hα sub-classification by brightness: F – faint, N – normal, B – bright
1 s.f.u. = 10^4 jansky = 10^{-2} W m^{-2} Hz^{-}

Vector Addition in Physics

Leonid Minkin[1], Alexander S. Shapovalov[2]

[1]*Department of Physics, Portland Community College, Portland OR, USA, www.pcc.edu/staff/lminkin*
[2]*Department of Physics, Saratov State University, Saratov, Russia*

(*Submitted* 20-01-2014)

Abstract

Introductory physics textbooks usually give some information about vectors and rules for their addition (triangle rule) and multiplication. In mathematics, addition of vectors is a matter of definition while in physics, in some cases, it must be done based on the meaning of addition and experimental confirmation. In physics, the triangle rule might be worthless if the meaning of addition is not strictly defined.

Keywords: vectors' addition, triangle rule, polar and axial vectors, sliding and bound vectors, center of boyuoncy, superposition

Introduction

Many fundamental principles of physics are presented in vector form. In mathematics, a vector is invariant with respect to rotation of coordinate axes and displacement from origin. Due to this property of vectors, the equations which express physics laws in vector form do not depend on the choice of inertial system of coordinates. This is the reason why it is simple and convenient to present physics laws in vector form. However, using vectors raises questions of transformation of different vector quantities under transition from one inertial (and non-inertial) system to another. The triangle rule is a mathematical rule for addition of two vectors. Nevertheless, this rule should be used with caution because in some cases, the physical meaning of vector addition must be first clarified.

In mechanics textbooks, the triangle rule of vector addition, including addition of velocities and accelerations, is usually described. This formal rule is applicable in Newtonian mechanics – for non-relativistic motion in inertial frames. On the contrary, in relativistic kinematics, where physical principles of measurements of length and time must be specified, the triangle rule of vector addition is not valid and the rule of vector addition should be based on physical principles of the special theory of relativity. In non-relativistic mechanics, the triangle rule for addition of accelerations in the frame rotating with respect to inertial one, in general, is not valid either.

The triangle rule is not always operational for addition of parallel vectors. For example, for addition of two collinear angular velocities, the location of instantaneous axis of rotation and magnitude of sum of angular velocities cannot be found based only on principles of vector algebra. Physical principles of kinematics must be used. Similar situation arises for addition of pressure forces applied to the object submerged in fluid. In this case, rules of vector algebra are not enough to find the point of application of the resultant of these forces – physical principles must be used.

Some examples, which illustrate the statements given in the abstract and introduction sections, are presented below.

Specifics of Vector Addition in Physics

1. Let us consider stationary frame K_1 and frame K, moving with the constant velocity $\boldsymbol{u}$ with respect to frame K_1. Imagine that the point object is moving with velocity $\boldsymbol{v}$ with respect to frame K. What is the velocity of the point object, $\boldsymbol{v_1}$, with respect to frame K_1? In this case, the motion of the object is considered in two different systems and clocks and rulers in different frames measure velocities. In classical mechanics, addition of velocities follows the triangle rule (mathematical vector addition) and $\boldsymbol{v_1} = \boldsymbol{v} + \boldsymbol{u}$, while in relativistic kinematics this rule cannot be applied ($\boldsymbol{v_1} \neq \boldsymbol{v} + \boldsymbol{u}$). Addition is based on the experimentally confirmed principles of special relativity theory. In one-dimensional case

$$v_1 = \frac{v+u}{1+\frac{vu}{c^2}} \qquad (1)$$

Here, c is the speed of light in vacuum.

2. In non-relativistic classical mechanics, addition of vectors of acceleration, in general, does not follow the triangle rule. Let the frame K be moving with acceleration $\boldsymbol{a}$ with respect to frame K_1. Imagine that a point like object is moving with acceleration a' with respect to frame K. In this case, the acceleration, $\boldsymbol{a_1}$, of the moving point in the system K_1 can be found by using the triangle rule only if both motions (K system and the point with respect to K frame) are translational. In this case $\boldsymbol{a_1} = \boldsymbol{a} + \boldsymbol{a}'$. In general,

$$\boldsymbol{a_1} \neq \boldsymbol{a} + \boldsymbol{a}'$$

and $\boldsymbol{a_1}$ is not defined by $\boldsymbol{a}$ and $\boldsymbol{a}'$ For example, if the coordinate system K is rotating with constant angular velocity $\boldsymbol{\omega}$ with respect to a fixed coordinate system K_1, the acceleration of the particle, $\boldsymbol{a_1}$, observed in the system K_1 is [1]

$$\boldsymbol{a_1} = \boldsymbol{a} + \boldsymbol{a_{cor}} + \boldsymbol{a}'$$

Here $\boldsymbol{a_{cor}} = \boldsymbol{2\omega} \times \boldsymbol{v_{rel}}$ is Carioles' acceleration, $\boldsymbol{v_{rel}}$ is the velocity of the point with respect to system K, $\boldsymbol{a}' = \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r})$ is the centripetal acceleration directed to the axis of rotation, and $\boldsymbol{r}$ is a vector position of the point in the system K.

3. In mathematics, the transformation of vector components is defined when one system of coordinate is turned with respect to the other. But the problem of transformation of vectors of electric and magnetic fields in moving systems of coordinates cannot be resolved using only mathematical principles. The transformation of electric and magnetic fields under

transition from one reference frame to another is not the same as transformation of vectors – electric and magnetic fields are coupled. Actually, electric and magnetic fields are second rank tensors [2] (although a vector can be considered as a first rank tensor). The special theory of relativity allows us to find the laws of conversion of electric and magnetic fields.

4. Superposition principle for electric and magnetic fields [2] is not obvious but is an experimental fact and it should be emphasized that this principle and the triangle rule of vector addition can be applied only to linear systems. Maxwell's equations, which are mathematical expressions of experimentally confirmed laws of nature, are linear and, therefore, the mathematical rule of vector addition is valid.

a. In mathematics, a directed line segment is called a vector [3]. Algebraically, a vector is a set of three numbers, which we call vector components [3]. These definitions are acceptable in physics although the given geometric and algebraic definitions suggest the concept of a sliding (free) vector. That is a vector whose initial point can be any point on a straight line that is parallel to the vector. In physics, statics, structures, and strength of materials, the point of vector application can be critical (bound vectors) [4] and the point of application of the resultant vector must be defined not only by mathematical rules but also by applying physics meanings. For example:

a. Any axial vector which is a cross product of two polar vectors and includes a vector position (torque, angular momentum) depends on the axis chosen and the point of force applied.

b. Deformation of solids depends on the point of force applied. It means that two parallel forces of equal magnitude (which are identical in mathematical meaning) cause different changes in the shape of an object.

c. The point of application of the resultant pressure force (center of boyuoncy, ***CB***) acting on the object submerged in fluid (buoyant force, $\boldsymbol{F_b}$) is important for studying rotational motion, equilibrium, and stability of the object in fluid. Typically, US introductory physics textbooks do not consider this topic. However, it may be difficult to predict ***CB*** intuitively. For example, let us consider a cylinder submerged in liquid (Fig. 1). The magnitude of the buoyant force acting upward on a partially or fully submerged object is equal to the weight of the liquid displaced by that object and ***CB*** point is the center of mass of the fluid displaced by the floating or submerged body [5]. The ***CB*** point for this particular case, which is the result of addition of parallel pressure forces, is located higher than all points of the appliedvertical pressure forces. The position of ***CB*** point is defined not by the mathematical rules of vector addition but by Newton's Principles of mechanics.

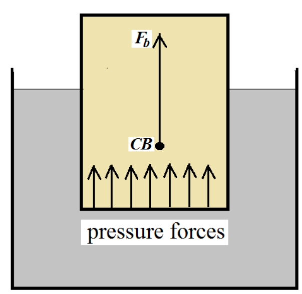


Fig 1: Vertical pressure forces, their resultant $\boldsymbol{F_b}$ (boyuant force), and center of boyuancy ***CB*** of the object submerged in liquid.

Similar comments can be made about the position of the center of gravity – the addition of gravity forces is defined not only by the mathematical rules of vector addition but by the definitions of center of gravity and torque.

d. Introductory physics textbooks typically consider smooth rolling objects (without slipping) along a flat surface. This motion can be studied as rotational motion with respect to the center of mass with angular velocity $\boldsymbol{\omega}$ and translational motion of the center of mass, or as pure rotational motion with the same angular velocity $\boldsymbol{\omega}$ with respect tomoving axis that always extends through contact lineof two surfaces [6]. However, a slightly different case of rolling thefirst cylinder of radius r_1 on the external surface of the second stationary cylinder of radius r_2 (Fig. 2) may be considered as pure rotation of the first cylinder with respect to its center (instantaneous axis *A*) with angular velocity $\boldsymbol{\omega}_1$ and pure rotational motion of the *A*-axis with respect to axis *B* with a *different* angular velocity $\boldsymbol{\omega}_2$. $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ are parallel vectors with respect to different frames (similar problem of rolling one cone along the surface of the other fixed one with both cones having common vertex is considered in [7] and [8]) and, according to the mathematical rule, their sum is

$$\boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2 \quad (2)$$

However, in this case, the application of the mathematical rule for vector additionis not enough to find the location of axis of rotation for which resultant motion of any point of the first cylinder is purely rotational with angular velocity $\boldsymbol{\omega}$**.** Eq. (2) does not define $\boldsymbol{\omega}_2$ or the magnitude of the resultant instantaneous angular velocity $\boldsymbol{\omega}$. To resolve above problems, additional kinematics relationships must be used.

Taking into account thatforrotation of the point on the rim of the first cylinder, the length of arc OO' is equal to the distance traveled in the same time by the point due to rotation with respect to axis *A* (Fig. 2b), one can see that for rolling without sliding

$$\boldsymbol{\omega}_2 = \boldsymbol{\omega}_1 \frac{r_1}{r_2} \quad (3)$$

Since the triangle rule foraddition of linear velocities can be applied in a non-relativistic approach, velocity $\boldsymbol{v}$ of any point ofthe first cylinder with respect to the second stationary cylinder is

$\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$

where $\boldsymbol{v}_1$ is velocity with respect tothe *A* axis and $\boldsymbol{v}_2$ is the velocity caused by rotation of the *A* axis with respect to axis *B*.For a point on the axis of rotation $\boldsymbol{v} = 0$, $\boldsymbol{v}_1 = -\boldsymbol{v}_2$ (Fig. 2c) can be accomplished only if the axis of rotation is a line of contact of the two cylinders (*O*-axis) for which Eq. (3) holds. It means that the first cylinder motion in frame *B*can be considered as pure rotation with angular velocity $\boldsymbol{\omega}$ with respect to the *O*-axis.

For the case of rolling of the first cylinder inside the hollow second cylinder (Fig. 2d), Eq. (11) holds, while angular velocities $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ are antiparallel and Eq. (3) must be replaced by

$$\boldsymbol{\omega}_2 = -\boldsymbol{\omega}_1 \frac{r_1}{r_2}$$

For both cases (rolling along internal or external surfaces), angular velocity with respect to instantaneous *O*-axis can be written in the form

$$\boldsymbol{\omega} = \boldsymbol{\omega}_1(1 \pm \frac{r_1}{r_2})$$

As r_2 approaches infinity (flat surface), the last equation transforms to a well known equality $\boldsymbol{\omega} = \boldsymbol{\omega}_1$.

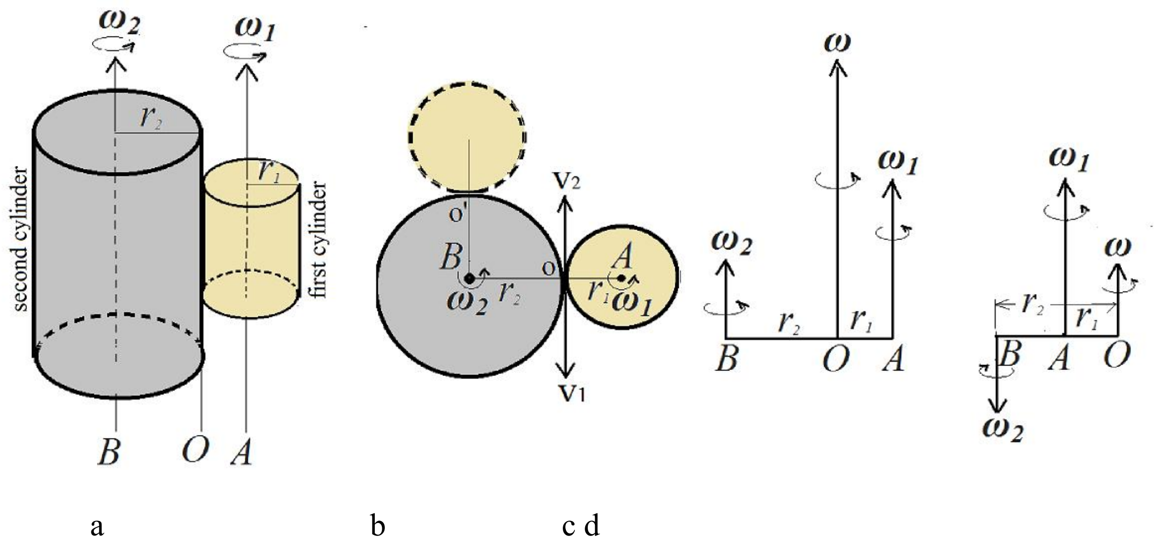


Fig 2: Smooth rolling of the first cylinder (radius r_1) on the external surface of the second stationary cylinder (a), its vertical projection (b) and vector diagram for addition of angular velocities (c); vector diagram of angular velocities for rolling of the first cylinder inside of the second hollow cylinder (d).

Conclusion

In mathematics, the rule of addition (triangle rule) of vectors is defined axiomatically, and conversion of vector components (three numbers) for transition from one Cartesian system of coordinates to another is defined by algebraic and trigonometric formulas. In physics, there are some occasions when operation of vector addition is based on its meaning and experimental justification.

It is obvious that physical restrictions on speeds cannot be explained by mathematics but special theory of relativity does set a limit for the speed of an object. Actually, all special relativity theory expressions are introduced with the use of a light signal exchange (the method of Einstein's synchronization). The light-signal method is used for time synchronization and for measuring length. This means that kinematics of the special theory of relativity is based on experiments (real or mental) which lead to rules of addition for velocities that are different from the triangle rule. Moreover, as a peculiarity of the special theory of relativity, the general law of velocity composition is not commutative [9] while the Galilean velocity composition law is commutative. Einstein's velocity addition is commutative only when **u** and **v** are parallel (Eq. (1)).

In some instances, the point of application of resultant force is important. However, vector algebra alone does not allow for locating of this point. For example, centers of gravity and buoyancy, which are the points of application of the forces of gravity and buoyancy, represent unique points of an object or system (the centers of gravity and buoyancy are not necessarily inside the object) which can be used to describe the system's response to external forces and torques. For example, the center of gravity of an extended body or system of masses is distinguished by the fact that it will remain at rest or moving at constant velocity unless the body is acted on by a net external force. The center of gravity may also be defined by implying that the torque about the origin would be the same if the entire weight acted through the center of gravity instead of acting through the individual masses. Similar rules should be used to find the center of bouncy. It means that principles of physics (Newton's Laws) must be exploited to locate the point of application of the resultant force. Another example, given in this paper, illustrates that to find instantaneous angular velocity (magnitude and location of axis of rotation) of a rolling cylinder on the surface of the second one, the vector addition rule of vector algebra is not enough – definitions of instantaneous axis of rotation and kinematics relationships are needed to be used. The transformation of electric and magnetic fields in moving systems of coordinates cannot be found using solely mathematical principles of vector operations. These transformations are not the same as transformation of vectors. When it comes to electric and magnetic field transformations, electric and magnetic fields are coupled. Electric and magnetic field vectors, as well as the vector of the electromagnetic force, are not invariant to different moving inertial frames, and the special theory of relativity allows us to find the laws of conversion of electric and magnetic fields and of Lorenz's force [2].

Typically, the considered specifics of vector addition are not the topic of introductory physics textbooks. It seems that some particularities of vector addition should be mentioned in the introductory physics class. Also, a brief review of polar and axial vectors, sliding and bond vectors would be helpful when students encounter problems similar to the ones presented here.

Acknowledgment

The authors want to thank professors W.Dittrich and R. Drosd for useful disscussions.

References

[1] A. P. French. *Newtonian Mechanics*, W. W. Norton (1971).

[2] L. D. Landau, E.M. Lifshitz. *Mechanics and Electrodynamics*, Pergamon Press(1972).

[3] E. Kreyszig. *Advanced Engineering Mathematics*, 10^{th} Ed., John Wiley&Sons, (2011).

[4] Tang, Kwong-Tin. *Mathematical Methodsfor Scientists and Engineers 2,*Springer(2007).

[5] F. M. White, *Fluid mechanics*, 7th edition, McGraw-Hill, New York(2011).

[6] D. Holliday, R. Resnik, J. Walker. *Fundamentals of Physics*, 10th Ed., John Wiley&Sons, Inc.(2013).

[7] Д.В. Сивухин, *ОбщийКурсФизики*. Т. 1. Механика, Физматгиз(2006).

[8] J. Peraire, S. Widnall <*http://ocw.mit.edu/courses/aeronautics-and-astronautics/16-07-dynamics-fall-2009/lecture-notes/MIT16_07F09_Lec25.pdf*>, visited in December 8, 2013.

[9] A. Ungar. The Relativistic Velocity Composition Paradox and the Tomas Rotation. *Foundation of Physics*, V. 19, #11, 1989.

The Requirement for Complex Numbers in Quantum Theory

G. Maynard[1] and D. Lambert[1] and W. D.Deering[1]

[1]Department of Physics
University of North Texas
Denton, TX USA.

(*Submitted*13-04-2015)

Abstract

Generations of physicists have at some point wondered about the role of complex numbers in quantum theory. None have had this point explained. Taking this as a pedagogical issue, this article elucidates the origins of the complex requirement and draws the connection between these origins and the requirement that the state vector be complex.

1. Introduction

A recent article by Sivakumar[1] highlights an important pedagogical void which has long plagued the teaching of introductory and fundamental principles of quantum theory. That is, there is no generally accepted explanation available for the apparent requirement for complex numbers in the mathematical formulation of the theory. Sivakumar provides a simplified demonstration (due to Sakurai[2] and Townsend [3]) that, in fact, the complex numbers are required. In this article, we continue beyond the demonstration and attempt to explain the underlying issues.

Quantum theory employs unit vectors to mathematically represent states of physical objects. In the following, we will identify four requirements which must be satisfied by these vectors. We then show that the four requirements are not satisfied by real vectors, but can be satisfied by complex vectors.

In the next section, we begin by simply stating the four requirements. We then devote an individual section to each requirement and discuss the physical and theoretical origins of that requirement.Turning from origins, we then consider them simply as a set of requirements on vector structure. By identifying this set of specific requirements on the vectors, we see clearly at exactly what point and for exactly what reason real vectors fail to satisfy the requirements.

We note up front that the article intent is pedagogical. Accordingly, points are presented in what is, hopefully, an intuitive, and conceptual way. We ask some leeway in completeness and rigor.

We also note that some closely related issues are put aside. For example, why the theory adopts use of vectors and the Born rule as a representational convention is an important foundational question. Here, we accept as a starting point that the theory does this. By making these choices, we focus on the question of, assuming that vectors are to be used in this way, why must the vectors be complex.

2. The Requirements

As mentioned, the theory uses vectors to represent states of physical objects. In particular, we will be interested in the representation of angular momentum states. It is in representing these states that the complex requirement arises.

The following four requirements must be satisfied by any vector, V, used by the theory to represent an angular momentum state. We will show that they are not satisfied by real vectors.

R1: Vector, V, is subject to n independent constraints with respect to an orthonormal basis set of vectors$\{ b_i \}_{(i = 1 , n)}$. Each constraint is of the form, $p_i = | (V , b_i) |^2$ (where the parenthesis indicates inner product).

R2: Vector V is n-dimensional.

R3: Vector V must vary with two real variables, "r" and "c".

R4: The set of constraints mentioned in R1 vary parametrically with the variable "r" above. That is, $p_i = p_i(r)$

In the next sections, we discuss these requirements individually with emphasis on their origins.

3. Requirement R1

In this section we state and discuss requirement R1.

R1: Vector, V, is subject to n independent constraints with respect to an orthonormal basis set of vectors $\{ b_i \}_{(i = 1 , n)}$. Each constraint is of the form, $p_i = |(V , b_i)|^2$.

Discussion:

In constructing the mathematical structure of any physical theory, some convention must be adopted for the representation of physical phenomena by mathematical structures. Quantum theory adopts, by postulate, the following representational convention:

P1: States of physical objects are represented by unit vectors, V.

P2: The probability for a transition between two states is represented by the "Born Rule". The "Born Rule" yields the probability as an inner product function on two vectors, V and b, which represent the two physical states involved in the transition, $p = | (V , b) |^2$.

The important point that we recognize in this section is that adoption of the Born Rule, in fact, imposes a constraint on vector V relative to vector b.

The familiar use of the Born Rule is to enter with the two state vectors, V and b, and obtain the transition probability, p. Here, we are recognizing a different perspective. It is the probability that is the observed physical fact. The vectors are merely mathematical structures employed to represent physical states. By adopting the Born convention to represent transition probabilities we are required to choose vectors which yield the correct probability value. From this perspective, the Born Rule, in fact, defines the vector pair (a partial definition) by specifying their relation. Consequently, we recognize a Born Rule expression as a "Born Constraint" on a state vector, V, relative to a transition state vector, b.

In addition to recognizing that the Born Rule imposes Born Constraints, requirement R1 also claims that there are n independent Born Constraints (with respect to the basis set). How do we know this?

A Born expression, $p = | (V , b) |^2$, represents the probability for a single transition from one state to another. It is observed physical fact, however, that an object in a given state can transition into one of some number, n, of alternative possible transition states. Each transition has some observed probability, p_i , and since they are mutually exclusive and exhaustive, $\Sigma_{(i = 1 , n)} (p_i) = 1$. This fact about probabilities imposes a requirement on the set of Born expressions representing the probabilities for the set of possible transitions.

That is:

$1 = \Sigma_{(i = 1 , n)} (p_i) = \Sigma_{(i = 1 , n)} (| (V , b_i) |^2)$, (Eqn. 1)

We can recognize this as, in fact, a requirement on the vector space used to represent states by the Born rule. That is, the vector space must come equipped with a defined L^2 vector norm.

The L^2 vector norm is defined as follows:

$| V | = \Sigma_{(i = 1 , n)} (| (V , b_i) |^2)$.

If vector V is a unit vector, then,

$1 = \Sigma_{(i = 1 , n)} (| (V , b_i) |^2)$. (Eqn. 2)

We see then that the choice to represent a set of transition probabilities by the Born Rule (Eqn. 1) has imposed the requirement that the vector space must be defined to have an L^2 vector norm (Eqn. 2).

Recognizing that the vector space has an L^2 vector norm is useful as follows.The set of vectors$\{ b_i \}_{(i = 1 , n)}$in (Eqn. 2)are an orthonormal basis set. Consequently there is a set of n individual Born Constraints on vector V, one associated with each basis vector. These constraints are independent because each is relative to a basis vector that is orthogonal to all of the others.

We therefore have the result R1 stated above. Any vector V used by the theory to represent the state of an object must satisfy requirement R1.

We note, in passing, an interesting and pedagogically useful point. The explanation of why quantum theory employs Hilbert space vectors to represent states is sometimes opaque. Here we understand that the theory makes an, early and fundamental commitment to the use of vector spaces which have an L^2 structure. If one generalizes the structure of a vector space in every way, dimensionality, etc., but retains the L^2 structure, then that is the set of Hilbert spaces. Quantum theory employs Hilbert spaces because the theory makes use of the L^2 structure.

4. Requirement R2

In this section we state and discuss requirement R2.

<u>R2: The vector is n dimensional</u>.

Discussion:

Having done the work of the previous section, we immediately recognize this requirement on any state vector. As explained, vector V is in a vector space spanned by the n orthonormal basis vectors $\{ b_i \}_{(i = 1 , n)}$. Consequently, V is n dimensional.

5. Requirement R3

In this section we state and discuss requirement R3.

<u>R3: Vector V must vary with two real variables, V (r, c</u>).

Discussion:

It is an observed physical fact that angular momentum states vary as a function of orientations or directions in physical space. The point is general, but can be seen by considering a simple example of two spin 1/2 objects. Suppose one object interacts with a Stern-Gerlach apparatus oriented in the z direction and deflects up along that direction. The second object interacts with a machine oriented along the (θ, ϕ) direction and deflects up along that direction. Subsequent to the interactions, these two objects are in objectively different physical states. What does it mean to be in different states? It means that subsequent observations made on the objects[1] will yield different results. They are observably different. We can state this same physical fact in another way by saying that angular momentum states vary with orientations in physical space.

It is a general point that in constructing a mathematical theory, for any mathematical object chosen to represent the physical state, that mathematical object must have the ability to vary as the physical state does. In particular, any vector we employ to represent angular momentum states must have the ability to vary with orientations in physical space. We can recognize this explicitly by writing the state vector as a function of orientation, V (O), where "O" is an orientation in physical space.

Orientations in three dimensional physical space vary with two degrees of freedom. Typically, polar coordinates, (θ, ϕ),are chosen to label spatial orientations. Here it will be useful to choose a different coordinatization. Select an arbitrary orientation, O_2 , then let real variables (r , c) label variation in radial and circumferential degrees of freedom relative to O_2 .

We can explicitly recognize this variation in two degrees of orientation freedom by writing the above state vector, V (O), as V (r , c) with "r" and "c" coordinates as defined.

We therefore have requirement R3 as given above. We note that the point here is to recognize that any vectors representing angular momentum states must have the ability to vary as the actual physical state varies, i.e., with two orientation degrees of freedom.

6. Requirement R4

In this section we state and discuss requirement R4.

<u>R4: The set of constraints mentioned in R1 vary parametrically with the variable "r" above. That is, $p_i = p_i(r)$</u>

(Since requirement R4 references the R1 Born Constraints, we copy again R1.

<u>R1: Vector, V, is subject to n independent constraints with respect to an orthonormal basis set of vectors $\{b_i\}_{(i = 1, n)}$. Each constraint is of the form, $p_i = |(V, b_i)|^2$</u>)

Discussion:

In the last section, we recognized the physical fact that angular momentum states vary with physical space orientations. Here, we recognize a second empirically observed fact characterizing angular momentum states. That is, for two angular momentum states associated with two different physical space orientations, O_1 , and O_2 , the probability for a transition from one state to the other varies as a function of the separation angle between the two orientations.

Here is where we can take advantage of the "r" and "c" coordinates defined earlier. If we take O_2 to be our arbitrary fixed reference, then the separation angle between the two orientations, O_1 , and O_2 , is given by the coordinate "r". Consequently, $p_i = p_i(r)$.

For the Born Constraints to vary parametrically with "r" we have made an assumption. That is, vector V is associated with one spatial orientation, O_1 , and all of the transition state vectors, $\{ b_i \}$ are associated with a

[1] The difference involves probabilities and consequently is observed on ensembles of similarly prepared objects.

single orientation, O_2 . This is appropriate for angular momentum observations. Suppose an object is in the state represented by vector, V (O_1). The object then interacts with a Stern-Gerlach apparatus oriented along O_2 . In this case, there are a set of n possible transition states, but we note the important fact that they are all associated with physical space orientation, O_2 .

Consequently, we have the result that the initial state vector is subject to a set of Born Constraints relative to the transition state vectors, { b_i }, and these constraints all vary parametrically with the separation angle parameter "r".

7. Satisfying The Requirements

In this section we collect again the four requirements and show that they are not satisfied by real vectors, but can be by complex vectors.

R1: Vector, V, is subject to n independent constraints with respect to an orthonormal basis set of vectors $\{ b_i \}_{(i = 1, n)}$. Each constraint is of the form, $p_i = |(V , b_i)|^2$

R2: Vector V is n dimensional.

R3: Vector V must vary with two real variables, "r" and "c".

R4: The set of constraints mentioned in R1 vary parametrically with the variable "r" above. That is, $p_i = p_i(r)$

We show first that real vectors do not satisfy these requirements as follows:

Point 1: Number of variables present

Assume V is real. Real vectors vary with one real variable per vector dimension. Requirement R2 requires that V is n-dimensional. Therefore, V varies with n variables.

Point 2: Number of constraints present

We see from R4 that the set of R1 constraints vary parametrically with variable "r". If we consider any fixed value of "r", then the R1 constraints impose n independent constraints on V (with respect to the orthonormal basis set { b_i }.

Points 1 and 2 imply: It follows that vector V is fully specified with respect to the basis set { b_i } (for any fixed "r"). There are n variables and n independent constraints. All variables present are assigned values by the constraints.

Therefore if vector V is real, and satisfies requirements, R1, R2, and R4, then it:

1. Is fully specified by "r", and
2. Varies as a function of "r".

Therefore:Having satisfied requirements, R1, R2, and R4, vector V cannot satisfy requirement R3. Vector V varies with and is fully specified by "r". Consequently, it is not possible for the vector to vary (nontrivially) in the second variable, "c", as is required by requirement R3.

We have shown that if vector V is real then it does not satisfy the set of requirements. Having done this analysis, however, one sees how substituting complex vectors for real vectors avoids the constraint limitation. The constraint encountered by real vectors is due to the availability of only n variables in the face of n constraints. A complex vector, however, provides 2n independent real variables. In the face of only n constraints, a 2n variable complex vector provides sufficient freedom to vary in both "r" and "c" degrees of freedom.

8. Discussion

Hopefully the analysis presented in this article is of pedagogical value. We have separated out issues in order to provide good access to the role of complex number in the theory.

We have identified a set of four specific requirements on the vectors employed by the theory. The goal was to facilitate two different perspectives linked by this set of requirements.

One perspective is mathematical. One can disregard the origins of the requirements and consider them simply as given. From this starting point, the exercise is one of vector structure. One can observe the interplay of freedom and constraint considerations that prevent the requirements from being satisfied by real vectors.

The other perspective is physical.Here we disregard the mathematical implications, and trace back the origins of the requirements. What specific features of the physical phenomena or adopted theoretical conventions impose these requirements?

The set of four requirements therefore serves as a point on which to stand and contemplate both available perspectives. From there, the student of foundational quantum theory can find a traceable connection all the way from the physical and theoretical origins through to their end consequence, a particular mathematical detail in the formal theory, the presence of complex numbers. More importantly, the student has a useful framework to separate out issues and make their own evaluation of the requirements, their origins, and their implications.

We point out two particular results of our analysis.

It is sometimes commented in the foundational literature that to explain some particular mathematical detail of the theory would be to elucidate its physical origins. Here we see that there is an identifiable physical origin. The complex requirement is, in part, a consequence of the fact that angular momentum states vary in two physical space orientation degrees of freedom. There is, however, a second equally important origin. It is the theory's adoption of a particular representational convention that imposes very substantial pairwise constraints on the vectors employed. Thus we see that, in this case, elucidating physical origins is not sufficient. We also must elucidate the theoretical representational conventions adopted.

We also mention a second result. We now have an answer to the big question, why are complex numbers required by the theory? The fundamental reason that they are required is to resolve a disappointingly mundane issue of freedom versus constraint. Quite simply, they provide more variables than real vectors. State vectors are subject to the significant "Born Constraints" yet must also honor a freedom demand when representing angular momentum states. The vectors must satisfy both. As we have seen, real vectors come up short on available variables. Consequently, we find complex vectors employed to represent states of physical objects.

References :

[1] S.Sivakumar, Physics Education, **28**No3.(2012)

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

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

Book Review:

SLEEPING BEAUTIES IN THEORETICAL PHYSICS

26 surprising insights

Thanu PADMANABHAN Springer 2015

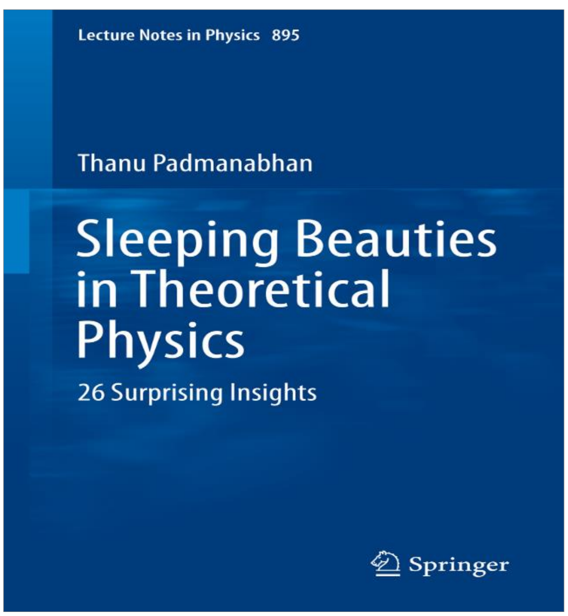
Lecture Notes in Physics: ISSN 1616-6361 (electronic)
ISBN 978-3-319-13442-0; ISBN 978-3-319-13443-7 (e-book) I-XVI + 302 pages

Reviewed by

R Ramachandran

Formerly
Institute of Mathematical Sciences, Chennai 600113
and
Indian Institute of Technology, Kanpur, Kanpur 208016

(rr_1940@yahoo.co.in)



Theoretical Physics** is Fun*: starts the preface to this collection of nuggets. I feel, that is about the apt definition to describe what is regarded as Theoretical Physics. Perhaps that it is fun justifies the pursuit of such a subject as an independent study more than any of its many accomplishments. Paddy, as Thanu Padmanabhan is known to many of us, brings together in this collection 26 pedagogic delights that will surely augment the understanding of many topics in Classical Mechanics, Quantum Mechanics, Special and General Relativity, Thermodynamics etc. covered in many traditional UG and PG programs in colleges and University. Some of these topics have earlier appeared in the sequence '**Curiosities in Theoretical Physics'** in Volumes 22(2005) - 24(2007) of this journal (*IAPT **Physics Education) as well as many more in several issues of ***Resonance*** (of *Indian Academy* of *Sciences*, Bangalore), as provided in the references in the book. I am delighted that they have now appeared in a well-connected, nevertheless self-contained, compilation.

Physics is driven by experiments and appears to encompass all of Natural Sciences. Dirac is supposed to have remarked that his equations for electrons explain most topics in Physics and all of Chemistry. Physics has thereafter had major influence on the analytical developments in Biology, particularly since the development of Molecular Biology and other hyphenated Life Science directed subjects. Most, if not all Engineering disciplines, derive their basics from Physics. Where does Theoretical Physics belong? Is it really a part of Applied Mathematics or is it more a vital corner of Physics, which in fact is clearly the experimental quest of the Nature. Does it depend on unambiguous experimental signals or is it an attempt to get an aesthetically pleasing viewpoint? As one plods through the 26 beauties with Padmanabhan, one cannot escape savouring the delightful part of the nature revealed by the tools of Theoretical Physics.

Everyone knows that planets move in ellipses – however, few know that they actually move in *circles* in the velocity space. Even for parabolic and hyperbolic trajectories in Kepler/Coulomb problem, the projectiles move in circular arcs in velocity space!

Inverse square force law is very special. Here Runge – Lenz vector is an additional conserved quantity, and is a cause of what is referred to as dynamical symmetry or *accidental symmetry*. In case you have been wondering what it specifies in an elliptical orbit, Paddy makes it clear that it is related to the vector that is directed from the centre of force, which is located at one of the two foci of the ellipse, to the other focus. If we choose to plot the position vector $\boldsymbol{r}$ and the velocity vector (rather a quantity with dimension of length, say, $\boldsymbol{v}' = L \times \boldsymbol{v}/|E|$), while $\boldsymbol{r}$ describes an ellipse, $\boldsymbol{v}'$ traces a circle about the second focus!

A spherically symmetric mass or charge distribution indeed causes a 1/r potential and consequent inverse square law outside. Is the converse true? Does inverse square law always imply spherically symmetric distribution of source? It is somewhat contrary to intuition that the answer is **no**! A simple example is from the method of images in Electrostatics. A conducting sphere has charges induced on the surface (clearly not a spherically symmetric distribution, indeed dipolar) that we know is *equivalent* to an image charge at a specific image point within the sphere for the force field in the region *outside* the sphere; and hence should lead to $1/r^2$ force field. Padmanabhan shows that even a planet shaped like a *deceased*(?) potato through some explicit inversion transformation can become one that will lead to $1/r$ type potential outside a defining sphere containing the weird shaped mass distribution.

Three body problem is notoriously hard. However there are tractable 3-body problems in a few special cases. If one of the three is light enough that it does not influence on the heavier pair, there are stable configuration for the threesome. To analyze them we may go to the co-moving rotating frame of the heavier pair and look for the potential extrema. There are 5 such points named after Lagrange and two of them are stable. It is surprising that the configuration is an equilateral triangle whatever be the actual masses (so long as one of them light as compared to other two) and the nature makes good use of it. At the Lagrange points L_4 and L_5 (which incidentally are local maxima of potential contours, nevertheless stable configuration with Coriolis forces providing the stability) of Sun – Jupiter system are located hundreds of Asteroids. They are referred to by Astronomers as Trojans and are given names after Greek (in L_4 region) and Trojan (in L_5 region) mythical heroes!

We know that classical mechanics is what one should get from Quantum Mechanics when the Planck constant goes to zero (and when the phase of the wave functions rapidly oscillate). How does one pick the classical endpoint? The variational principle that we use in classical Physics emerges naturally. It is remarkable, as the example clarifies, that we can better understand aspects of classical mechanics when we approach it from the quantum regime!

I am surprised to learn that even for describing a free particle motion, rather for defining the relevant Lagrangian, fully relativistic form is logically better than the nonrelativistic limit. Likewise Klein Gordon equation is more logical than the Schroedinger equation that it leads to in the non-relativistic limit.

Several text books show how to solve for quantum levels for Simple Harmonic motion both by means of a differential equation and by algebraic methods using ladder operators. Hydrogen atom states can also be obtained by algebraic means – though this is known to many, most textbooks have not exposed how this is done. What is involved is no more than what one knows from angular momentum algebra. It also happens to be the best way to drive home the fact that the energy levels depend only on the principal quantum number (independent of *l* values).

It is indeed curious to know that the Hydrogen problem can be paraphrased as harmonic oscillator in four dimensions with some appropriate constraints. That means the only problem that we can solve exactly in Quantum Mechanics (apart from the free particle in a box) is the simple harmonic motion and (the myriad forms it assumes on each occasion, including Hydrogen atom as we find now and indeed the theory of Strings).

Astrophysicist S Chandrashekhar described the Black Hole as a perfect end state – a state believed to be characterized by just Mass, Angular Momentum and Charge. When they acquired entropy and temperature without any clue on its internal structure and degrees of freedom associated with them, it was puzzling. Quantum Mechanics comes to the rescue with Hawking radiation. Is there a simple way to attribute Temperature to a Black Hole?

Probability plays a significant role in the nature of physical phenomena. We first encounter it in the topics of Heat and Thermodynamics. The Second law of thermodynamics introduces a measure of randomness as an extensive variable called Entropy. Maximizing entropy is the way to go, says the Second Law of Thermodynamics. Quantum mechanics are forever about probability amplitudes. Is there a connection? Feynman path integrals (for time evolution of the states) are closely related to Wiener integrals in probability summations of equilibrium systems with the temperature $1/\beta$ playing the role of imaginary time. In this context Random walks are delightful tools to analyse the many natural phenomena. There are two beauties in this collection for you to marvel.

The collection is a highly suitable pedagogic material and should be a **must read** for anyone who is interested in Physics, either for a career or for pleasure. It should require no more expertise than is available to an advanced undergraduate student of Physics or allied disciplines. I recommend that this is definitely acquired by all college libraries.

Finally why 26? Is there any numerical significance? I suppose it does not have anything to do with the number of required dimensions for a String theory with bosonic co-ordinates?

******

Excerpts from the Book ….

Chapter Highlights:

1. **The Grand Cube of Theoretical Physics**

The 'big picture' of theoretical physics can be nicely summarized in terms of a unit cube made of the fundamental constants G, h, c^{-1} representing the three axes. The vertices and linkages of this cube — which we will explore in different chapters of this book — allow you to appreciate different phenomena and their inter-relationships. This chapter introduces the Cube of Theoretical Physics and relates it to the rest of the book.

2. **The Emergence of Classical Physics**

Quantum physics works with probability amplitudes while classical physics assumes deterministic evolution for the dynamical variables. For example, in non-relativistic quantum mechanics, you will solve the Schrodinger equation in a potential to obtain the wave function $\psi(t,q)$, while the same problem — when solved classically will lead to a trajectory $q(t)$. How does a deterministic trajectory arise from the foggy world of quantum uncertainty?

We will explore several aspects of this correspondence in this chapter, some of which are nontrivial. You will discover the real meaning of the Hamilton-Jacobi equation (without the usual canonical transformations, generating functions and other mumbo-umbo) and understand why the Hamilton-Jacobi equation told us $p_\alpha = \partial_\alpha S = (-\partial_t S, \nabla S) = (E, p_x\ p_y\ p_z)$ even before the days of four-vectors and special relativity. We will also address the question of why the Lagrangian is equal to kinetic energy minus potential energy (or is it, really?) and why there are only two classical fields, electromagnetism and gravity. In fact, you will see that classical physics makes better sense as a limit of quantum physics!

3. **Orbits of Planets are Circles!**[1]

The orbits of planets, or any other body moving under an inverse square law force, can be understood in a simple manner using the idea of the velocity space. Surprisingly, a particle moving in an ellipse, parabola or a hyperbola in real space moves in a circle in the velocity space. This approach allows you to solve the Kepler problem in just two steps! We will also explore the peculiar symmetry of the Lagrangian that leads to the conservation of the Runge-Lenz vector and the geometrical insights that it provides. Proceeding to the relativistic versions of Kepler/Coulomb problem you will discover why the forces must be velocity dependent in a relativistic theory and describe a new feature in the special relativistic Coulomb problem, viz. the existence of orbits spiraling to the center.

4. **The Importance of being Inverse-square**

This chapter continues the exploration started in the previous one. The Coulomb problem, which corresponds to motion in a potential that varies as r^{-1}, has a peculiar symmetry which leads to a phenomenon known as 'accidental' degeneracy. This feature exists both in the classical and quantum domains and allows some interesting, alternative ways to understand, e.g., the hydrogen atom spectrum. We will see how one can find the energy levels of the hydrogen atom without solving the Schrodinger equation and how to map the 3D Coulomb problem to a 4D harmonic oscillator problem. The $(1/r)$ nature of the potential also introduces several peculiarities in the Scattering problem and we will investigate the questions:

(i) How come quantum Coulomb scattering leads exactly to the Rutherford formula? What happened to the $\hbar$?

(ii) How come the Born approximation gives the exact result for the Coulomb potential? What do the 'unBorn' terms contribute?!

5. **Potential surprises in Newtonian Gravity**[2]

How unique is the distribution of matter which will produce a given Newtonian gravitational field in a region of space? For example, can a non-spherical distribution of matter produce a strictly inverse square force outside the source? Can a non-planar distribution of matter produce a strictly constant gravitational force in some region? We discuss the rather surprising answers to these questions in this chapter. It turns out that the relation between the density distribution and the gravitational force is far from what one would have naively imagined from the textbook examples.

6. **Lagrange and his Points**

A solution to the 3-body problem in gravity, due to Lagrange, has several remarkable features. In

[1] Curiosities in Theoretical Physics: IV IAPT *Phys Edu* **23** No. 3 Art. no. 7(2006 -07)

[2] Curiosities in Theoretical Physics: I IAPT *Phys. Edu.* **22** No. 4 art.no. 2(2005-06)

particular, it describes a situation in which a particle, located at the maxima of a potential, remains stable against small perturbations. We will learn a simple way of obtaining this equilateral solution to the three body problem and understanding its stability.

7. **Getting the most of it!**

Extremum principles play a central role in theoretical physics in many guises. We will discuss, in this chapter, some curious features associated with a few unusual variational problems. We start with a simple way to solve the standard brachistochrone problem and address the question: How come the cycloid solves all the chronic problems? (Or does it, really?). We then consider the brachistochrone problem in a real, ($1/r^2$), gravitational field and describe a new feature which arises: viz. the existence of a forbidden zone in space not accessible to brachistochrone curves! We will also determine the shape of a planet that exerts the maximum possible gravitational force at a point on its surface — a shape which does not even have a name! Finally, we take up the formation of the rainbows with special emphasis on the question: Where do you look for the tertiary (3rd order) rainbow?

8. **Surprises in Fluid Flows**

The idealized flow of a fluid around a body is a classic text book problem in fluid mechanics. Interestingly enough, it leads to some curious twists and conceptual conundrums. In particular, it leads a surprising divergence which needs to be regularized even in the text book case of fluid flow past a sphere!

9. **Isochronous Curiosities: Classical and Quantum**

The oscillatory motion of a particle in a one dimensional potential belongs to a class of exactly solvable problems in classical mechanics. This chapter examines some lesser known aspects of this problem in classical and quantum mechanics. It turns out that both $V(x) = ax^2$ and $V(x) = ax^2 + bx^{-2}$ have (1) periods of oscillation which are independent of amplitude in classical physics and (2) equally spaced energy levels in quantum theory. We will explore several features of this curious correspondence. We will also discuss the question of determining the potential from the period of oscillation (in classical physics) or from the energy levels (in quantum physics) which are closely related and clarify several puzzling features related to this issue.

10. **Logarithms of Nature**

Scaling arguments and dimensional analysis are powerful tools in physics which help you to solve several interesting problems. And when the scaling arguments fail, as in the examples discussed in this chapter, we are led to a more fascinating situation. A simple example in electrostatics leads to infinities in the Poisson equation and we get a finite E from an infinite ϕ! I also describe the quantum energy levels in the delta function potentials and show how QFT helps you to understand QM better!

11. **Curved Spacetime for pedestrians**[3]

The spacetime around a spherical body plays a key role in general relativity and is used in the crucial tests of Einstein's theory of gravity. This spacetime geometry is usually obtained by solving Einstein's equations. I will show how this metric can be obtained by a simple

[3] Curiosities in Theoretical Physics: II IAPT *Phys. Edu.* **23** 47(2006-07)

— but strange — trick. Along the way, you will also learn a three-step proof as to why gravity must be geometry, the reason why the Lagrangian for a particle in a Newtonian gravitational field is kinetic energy minus potential energy and how to obtain the orbit equation in GR, just from the principle of equivalence.

12. **Black hole is a Hot Topic**[4]

A fascinating result in black hole physics is that they are not really black! They glow as though they have a surface temperature which arises due to purely quantum effects. I will provide a simple derivation of this hot result based on the interpretation of a plane wave by different observers.

13. **Thomas and his Precession**[5]

Thomas precession is a curious effect in special relativity which is purely kinematical in origin. But it illustrates some important features of the Lorentz transformation and possesses a beautiful geometric interpretation. We will explore the physical reason for Thomas precession and its geometrical meaning in this chapter and in the next.

14. **When Thomas met Foucault**

The Foucault pendulum is an elegant device that demonstrates the rotation of the Earth. We describe a paradox related to the Foucault pendulum and provide a geometrical approach to determine the rotation of the plane of the pendulum. By introducing a natural metric in the velocity space we obtain an interesting geometrical relationship between the dynamics of the Foucault pendulum and the Thomas precession discussed in the previous chapter. This approach helps us to understand both phenomena better.

15. **The One-body Problem**

You might have thought that the one-body problem in physics is trivial. Far from it! One can look at the free particle in an inertial or a non-inertial frame, relativistically or non-relativistically, in flat or in curved spacetime, classically or quantum mechanically. All these bring in curious correspondences in which the more exact theory provides valuable insights about the approximate description. I start with the surprising — and not widely appreciated — result that you really can't get a sensible free-particle Lagrangian in non-relativistic mechanics while you can do it in relativistic mechanics. In a similar vein, the solution to the Klein-Gordon equation transforms as a scalar under coordinate transformations, while the solution to the Schrodinger equation does not. These conundrums show that classical mechanics makes more sense as a limiting case of special relativity and the nonrelativistic Schrodinger equation is simpler to understand as a limiting case of the relativistic Klein-Gordon equation!

16. **The Straight and Narrow Path of Waves**

Discovering unexpected connections between completely different phenomena is always a delight in physics. In this chapter and the next, we will look at one such connection between two unlikely phenomena: propagation of light and the path integral approach to quantum Field theory! This chapter introduces the notion of paraxial optics in which we throw away half the solutions and still get useful results! I also describe the role of optical systems and how the humble lens acts as an analog device that performs Fourier

[4] Curiosities in Theoretical Physics: V IAPT *Phys. Edu.* **24,** 67(2007-08)

[5] Curiosities in Theoretical Physics: VII IAPT *Phys Edu* **24**, 149(2007-08)

transforms. In passing, you will also learn how Faraday's law leads to diffraction of light.

17. **If Quantum Mechanics is the Paraxial Optics, then ...**

The quantum mechanical amplitude for a particle to propagate from event to event in spacetime shows some nice similarities with the corresponding propagator for the electromagnetic wave amplitude discussed in the previous chapter. This raises the question: If quantum mechanics is paraxial optics, what is the exact theory you get when you go beyond the paraxial approximation? In the path integral approach to quantum mechanics you purposely avoid summing over all the paths while in the path integral approach for a relativistic particle you are forced to sum over all paths. This fact, along with the paraxial optics analogy, provides an interesting insight into the transition from quantum field theory to quantum mechanics and vice versa! I also describe why combining the principles of relativity and quantum theory demands a description in terms of fields.

18. **Make it Complex to Simplify**

Some of the curious effects in quantum theory and statistical mechanics can be interpreted by analytically continuing the time coordinate to purely imaginary values. We explore some of these issues in this chapter. In quantum mechanics, this allows us to determine the properties of ground state from an approximate evaluation of path integrals. In statistical mechanics this leads to an unexpected connection between periodicity in imaginary time and temperature. The power of this approach can be appreciated by the fact that one can derive the black hole temperature in just a couple of steps using this procedure. Another application of the imaginary time method is to understand phenomena like the Schwinger effect which describes the popping out of particles from the vacuum. Finally, I describe a non-perturbative result in quantum mechanics, called the over-the-barrier-reflection, which is easier to understand using complex paths.

19. **Nothing matters a lot**

The vacuum state of the electromagnetic field is far from trivial. Amongst other things, it can exert forces that are measurable in the lab. This curious phenomenon, known as the Casimir effect, is still not completely understood. I describe how the probability distribution for the existence of electromagnetic fields in the vacuum can be understood, just from the knowledge of the quantum mechanics of the harmonic oscillator. This chapter also introduces you to the tricks of the trade in quantum field theory, which are essential to get finite answers from divergent expressions - like to prove that the sum of all positive integers is a negative fraction!

20. **Radiation: Caterpillar becomes Butterfly**

The fact that an accelerated charge will radiate energy is considered an elementary textbook result in electromagnetism. Nevertheless, this process of radiation (and its reaction on the charged particle) raises several conundrums about which technical papers are written even today. In this chapter, we will try to understand how the caterpillar ($1/r^2$, radial field) becomes a butterfly ($1/r$, transverse field) in a simple, yet completely rigorous, manner without the Lienard-Wiechert potentials or other red-herrings. I will also discuss some misconceptions about the validity of $\nabla\cdot E = 4\pi\rho$ for radiative fields with retardation effects.

21. **Photon: Wave and/or Particle**

The interaction of charged particles with blackbody radiation is of considerable practical and theoretical importance. Practically, it occurs in several astrophysical scenarios. Theoretically, it illustrates nicely the fact that one can think of the radiation either as a bunch of photons or as electromagnetic waves and still obtain the same results. We shall highlight some non-trivial aspects of this correspondence in this chapter. In particular we will see how the blackbody radiation leads a double life of being either photons or waves and how the radiative transfer between charged particles and black body radiation can be derived just from a Taylor series expansion (and a little trick)! Finally, I will describe the role of radiation reaction force on charged particles to understand some of these results.

22. **Angular Momentum without Rotation**[6]

Not only mechanical systems, but also electromagnetic fields carry energy and momentum. What is not immediately apparent is that certain static electromagnetic configurations (with no rotation in sight) can also have angular momentum. This leads to surprises when this angular momentum is transferred to the more tangible rotational motion of charged particles coupled to the electromagnetic fields. A simple example described in this chapter illustrates, among other things, how an observable effect arises from the unobservable vector potential and why we can be cavalier about gauge invariance in some circumstances.

23. **Ubiquitous Random Walk**

What is common to the spread of mosquitoes, sound waves and the flow of money? They all can be modelled in terms of random walks! Few processes in nature are as ubiquitous as the random walk which combines extraordinary simplicity of concept with considerable complexity in the final result. In this and the next chapter, we shall examine several features of this remarkable phenomenon. In particular, I will describe the random walk in the velocity space for a system of gravitating particles. The diffusion in velocity space can't go on and on — unlike that in real space — which leads to another interesting effect known as dynamical friction — first derived by Landau in an elegant manner.

24. **More on Random Walks: Circuits and a Tired Drunkard**

We continue our exploration of random walks in this chapter with some more curious results. We discuss the dimension dependence of some of the features of the random walk (e.g., why a drunken man will eventually come home but a drunken bird may not!), describe a curious connection between the random walk and electrical networks (which includes some problems you can't solve by being clever) and finally discuss some remarkable features of the random walk with decreasing step-length, which is still not completely understood and leads to Cantor sets, singularities and the golden ratio — in places where you don't expect to see them.

25. **Gravitational Instability of the Isothermal Sphere**

The statistical mechanics of a system of particles interacting through gravity leads to several counter-intuitive features. We explore one of them, called Antonov instability, in this chapter. I describe why the thermodynamics of gravitating systems is non-trivial and how to obtain the mean-field description of such a system. This leads to a self-gravitating distribution of mass called the isothermal sphere which exhibits curious features both from the mathematical and physical points of view. I provide a simple way of understanding the stability of this system, which is of astrophysical significance.

[6] Curiosities in Theoretical Physics: V IAPT *Phys. Edu.* **23** 285(2006-07)

26. **Gravity bends electric field lines**

Field lines of a point charge are like radially outgoing light rays from a source. You know that the path of light is bent by gravity; do electric field lines also bend in a gravitational field? Indeed they do, and — in the simplest context of a constant gravitational field — both are bent in the same way. Moreover, both form arcs of circles! The Coulomb potential in a weak gravitational field can be expressed in a form which has unexpected elegance. The analysis leads to a fresh insight about electromagnetic radiation as arising from the weight of electrostatic energy in the rest frame of the charged particle, and also allows you to obtain Dirac's formula for the radiation reaction in three simple steps

*****