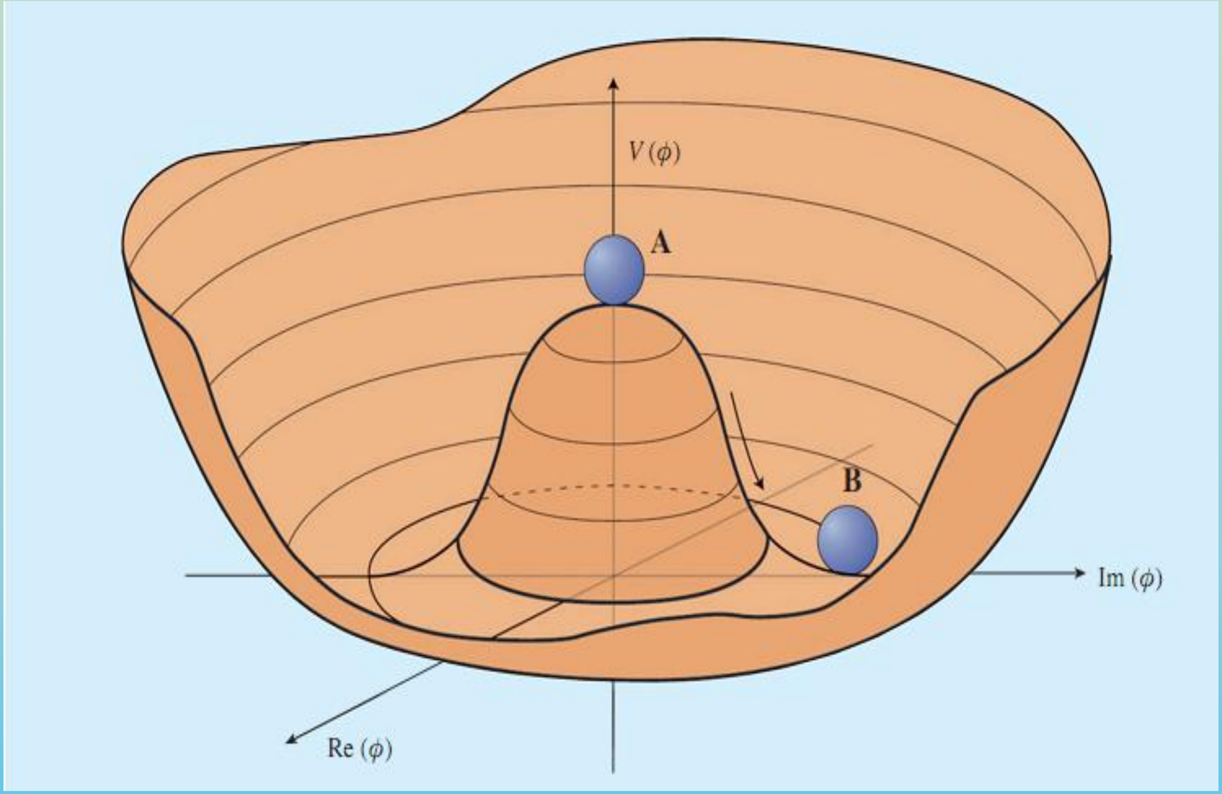
Vol 29 No. 3 Jul - Sep 2013 ISSN 0970-5953

PHYSICS EDUCATION



The variation of Higgs field with potential energy.

www.physedu.in

Volume 29, Number 3

In this Issue

EDITORIAL

It is a pleasure to publish number 29.3, although a bit late. This issue is a blend of articles on physics pedagogy and research in physics education. I must mention two articles on functional differential equations (FDES) by Prof. C. K. Raju. FDEs have a promise to give rise to a new formulation of dynamics, both classical and quantum. Professor Raju had written a series of articles in Physics Education on the concept of time in science, including his own original contributions. These articles are published as a book by Kluwer who are selling it at something like two hundred dollars! We hope that these articles on FDEs also develop into a series and readers get an opportunity to read them, before they start selling at some prohibitive price!

Next, the article by S. Ganci makes an interesting account of the systematic perturbation introduced by the measuring instruments. The article on three-level lasing system by Madhuramurthy and Sarmistha Sahu uses dice throwing as a pedagogical tool to understand various characteristics of three level lasing systems and makes an interesting reading. The articles by D. Banerjee and S. Sahoo and by Tapashi Das and D. K. Choudhury give pedagogical view of models of elementary particles and their interactions. Kamo I. Atayan discusses an important issue of incorporating environmental education as a part of physics teaching. This issue has a deep social significance. The article by Arunava Roy discusses the role of spectroscopy in understanding the HRD diagrams used to study properties of stars including our sun. This article has pointers useful in teaching these topics. The paper by N. Doumanis and Ch. Trikalinos shows effective ways of using Young's experiment as a pedagogical tool.

I wish you a very happy reading experience!

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

Functional differential equations. 1: A new paradigm in physics

C. K. Raju

AlBukhary International University
05200 Alor Setar, Malaysia
ckr@ckraju.net

(Submitted 22-07-2013)

Abstract

This article explains the basics of functional differential equations (FDEs). FDEs differ fundamentally from the ordinary and partial differential equations. As such, FDEs present a new paradigm in physics. Since FDEs arise naturally through Maxwellian electrodynamics, this involves no new physical hypothesis, but only a revised mathematical understanding of existing physics.

1 Introduction

1.1 ODEs and PDEs

Physicists have long been familiar with ordinary differential equations (ODEs). Newton's second "law" of motion reduces physical problems of classical mechanics to the mathematical problem of solving ODEs.

For example, as every physics student knows, a force proportional and opposite to displacement leads to the ODE of simple harmonic motion:

$$y'' = -k^2 y. \tag{1}$$

Here, $y \equiv y(t)$ denotes the position of the particle at time t, primes denote derivatives and k is a constant.

The second order ODE (1) has a unique solution if we prescribe the initial value of y and its first derivative y'.

$$\begin{aligned} y(0) &= y_0, \\ y'(0) &= y_1. \end{aligned} \tag{2}$$

In general, ODEs have a unique solution from (appropriate) initial data. Physically, this feature of ODEs means that the future states of a classical system are determined if we prescribe its initial state: namely the initial position and velocity.

In general, if we use Hamilton's form of the ODEs for a classical system of n particles, to fix the initial state we must prescribe the initial values of the positions and canonically conjugate momenta for each particle. This is the Newtonian paradigm: *the future states of a physical system are determined by its initial state alone.*

Less easy, but common in physics, are partial differential equations (PDEs). In classical physics, these arise when, instead of a system of n discrete particles, one works with continua—that is, fluids or fields. Some well-known PDEs are the Navier-Stokes equations for fluid flow, Maxwell's equations for electromagnetic fields, and the Hilbert-Einstein equations of general relativity. In quantum physics, even a single particle is described by PDEs such as the Schrödinger equation.

1.2 FDEs

Little known, however, are the functional differential equations (FDEs) which differ from both ODEs and PDEs. The mathematical theory of FDEs has been around for some time now, but their fundamental importance for physics was understood only relatively recently. Hence, these are not yet common in physics texts.

Let us start with a very simple example of an FDE. Consider

$$y'(t) = y(t-1). \tag{3}$$

This is also called a retarded FDE or a delay differential equation, since it involves a time-lag (retardation) or delay: instead of $y(t)$ we have $y(t-1)$ on the right hand side. What difference does that make? To see this, suppose we try to solve this FDE like we solve ODEs. That is, we give the initial value $y(0)$, and ask for the value of $y(1)$. Can we obtain it?

1.3 Retarded FDEs need past data

Symbolically we can write:

$$\begin{aligned} y(1) &= y(0) + \int_0^1 y'(t)dt && (4)\\ &= y(0) + \int_0^1 y(t-1)dt. && (5)\end{aligned}$$

But to actually carry out the integration we need to know the values of the integrand $y(t-1)$ in (5). To know $y(t-1)$ for all $t \in [0,1]$ is to know $y(t)$ for all $t \in [-1,\ 0]$. That is, unlike the ODE $y' = y$ for which the *initial* data at $t = 0$ suffices, for the retarded FDE (3) we need *past* data or the values of the function $y(t) \equiv \phi(t)$ over an entire past interval $[-1,\ 0]$.

If, instead of an initial function $\phi(t)$, only the initial value $y(0)$ is prescribed, then we can assume the past values in infinitely many different ways, so we can find an infinity of different solutions. Three such solutions are depicted in the graph: all these solutions have the same prescribed initial value of $y(0)$ but different past values (see Fig. 1).

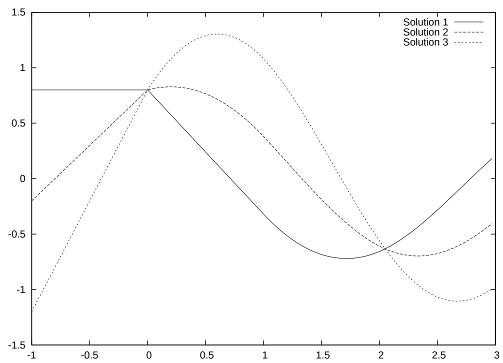


Figure 1: **FDEs need past data:** Three different solutions of the FDE (3) with the same initial data but different past data.

The equation (3) relates the rate of change of y *now* to its *past* values, i.e., the FDE models a history-dependent situation. So, intuitively, it makes sense that to solve it we need to know the relevant past data or past history. It is obvious from (5) that the FDE (3) has a unique solution if the initial function $\phi(t)$ is continuous. (However, this solution may fail to be differentiable at $t = 0$, as in Solution 1 of Fig. 1.) This can be generalised into a formal mathematical theorem[1] that FDEs admit a unique local solution from past data under appropriate conditions.

1.4 Time asymmetry of retarded FDEs

FDEs differ from ODEs in another fundamental way. ODEs are instantaneous: Newton's second "law" relates force *now* to acceleration (or the second derivative of position) *now*. Since ODEs relate to what happens at *one* instant of time, they do not discriminate between past and future. Instead of regarding (2) as "initial" data, we can just as well regard it as "final" data and solve (1) backward in time.

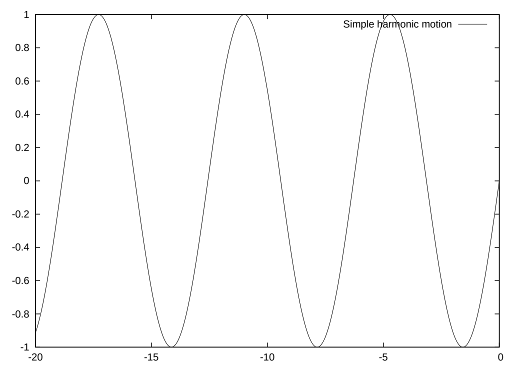


Figure 2: **An ODE solved backward.** ODEs are time symmetric; they can be solved either forward or backward in time. The above graph shows a solution of the simple harmonic oscillator calculated backward from prescribed values of $y(0)$ and $y'(0)$ regarded as final data. This can be done exactly as we calculate forward solutions with the same values regarded as initial data. The solution was obtained with my software CALCODE which solves ODEs in either direction in time.

The retarded FDE (3), however, is time *asymmetric*: it relates $y'(t)$ (or the rate of change of y now) to its past values, namely $y(t-1)$. It can only be solved forward in time. We can determine $y(t)$ for future values ($t \geq 0$) given past data ($y(t)$ for $t \leq 0$), but not the other way round: past values ($y(t)$ for $t \leq 0$) cannot be determined from future data ($y(t)$ for $t \geq 0$). In the following example, different past histories all converge to the same future, hence from a knowledge of the future, one cannot ascertain the past.

1.5 An example

Consider the FDE

$$y'(t) = b(t)y(t-1), \tag{6}$$

where b is a continuous function which vanishes outside [0, 1], and satisfies

$$\int_{-\infty}^{\infty} b(t)\ dt = \int_0^1 b(t)\ dt = -1. \tag{7}$$

For example,

$$b(t) = \begin{cases} 0 & t \leq 0, \\ -1 + \cos 2\pi t & 0 \leq t \leq 1, \\ 0 & t \geq 1. \end{cases} \tag{8}$$

For $t \leq 0$, the FDE (6) reduces to the ODE $y'(t) = 0$, so that, for $t \leq 0$, $y(t) = k$ for some constant k ($= y(0)$).

Now, for $t \in [0,\ 1]$,

$$\begin{aligned} y(t) &= y(0) + \int_0^t y'(s)ds \\ &= y(0) + \int_0^t b(s)y(s-1)ds \\ &= y(0) + y(0)\int_0^t b(s)ds, \end{aligned} \tag{9}$$

since $y(s-1) \equiv k = y(0)$ for $s \in [0, 1]$. Hence, using (7), $y(1) = 0$, no matter what k was. However, since $b(t) = 0$ for $t \geq 1$, the FDE (6) again reduces to the ODE $y'(t) = 0$, for $t \geq 1$, so that $y(1) = 0$ implies $y(t) = 0$ for all $t \geq 1$.

Hence, the past of a system modeled by (6) cannot be retrodicted from a knowledge of the entire future; for if the future data (i.e., values of the function for *all* future times $t \geq 1$) are prescribed using a function ϕ that is different from 0 on $[1,\ \infty]$, then (6) admits no backward solutions for $t \leq 1$. If, on the other hand, $\phi \equiv 0$ on $[1,\ \infty]$, then there are an infinity of distinct backward solutions. Fig. 3 shows three such solutions. In either case, knowledge of the entire future furnishes no information about the past.

With retarded FDEs we can infer future from past, but not, in general, past from future. They model a situation where we have *more* information about the past than the future. To put matters differently, we can say there is loss of information towards the future. This is the same as saying that retarded FDEs model a situation where there is an increase of entropy towards the future, for information is the negative of entropy, as I have explained in detail, in an earlier paper.[2] With ODEs,

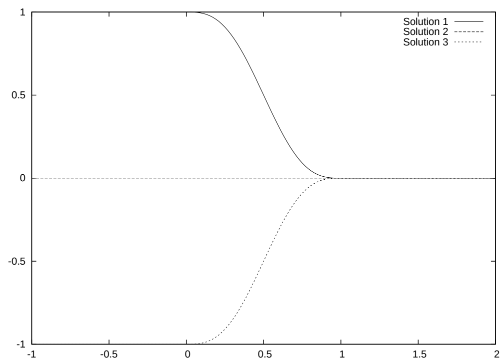


Figure 3: **FDEs are time asymmetric:** Three different solutions of a retarded FDE with the different pasts have the same future, so that retarded FDEs cannot be solved backward. That is, past cannot be inferred from knowledge of future.

entropy must stay constant.

2 The paradigm shift

These features—the need for past data, and time asymmetry—mean that the physics of a system modeled by FDEs differs fundamentally from the physics of a system modeled by ODEs (classical mechanics). This led to my claim that FDEs involve a shift away from the Newtonian paradigm.

A "paradigm shift" refers to a fundamental new idea. People often resist such new ideas, and try to hang on to the old ways of thinking. More specifically, physicists accustomed to ODEs and PDEs resist the changes necessitated by FDEs, and tend to fall back into the old ways associated with ODEs and PDEs.

2.1 Converting FDEs to ODEs is a mistake

This has led to a common mistake: physicists (including Einstein) tried to approximate FDEs by ODEs. Such an approximation seems plausible because the retarded FDEs which arise in physics typically involve "small" delays. That is, we have FDEs of the type

$$y'(t) = y(t - \tau) \tag{10}$$

where $\tau > 0$ is "small". Suppose we "Taylor" expand $y(t - \tau)$ in powers of the delay τ. This allows us to approximate $y(t - \tau)$ by the values of y and its derivatives at t. Inserting this approximation into the FDE (10) converts it to a higher-order ODE. Is this a valid approximation? Will the solutions of the resulting ODE approximate the solutions of the original FDE?

No. Not, in general. This should be obvious by now, because of the qualitative differences between FDEs and ODEs brought out above. However, let us see yet another explicit counter-example.

2.2 Incorrectness of approximating FDEs by ODEs

Consider the equation

$$y'(t) = y(t - \tau) - y(t), \tag{11}$$

where $\tau > 0$ is a small constant. If we expand the right hand side in "Taylor's" series in powers of the delay τ, and truncate after two terms, we obtain

$$y'(t) = \{y(t) - \tau y'(t) + \frac{\tau^2}{2!}y''(t)\} - y(t). \quad (12)$$

This simplifies to

$$y''(t) - \frac{2(1+\tau)}{\tau^2}y'(t) = 0. \quad (13)$$

This is a linear ODE with constant coefficients, which every physics student should know how to solve. The solution is

$$y_{\text{approx}} = c_1 + c_2 e^{st}, \quad (14)$$

where $s = \frac{2(1+\tau)}{\tau^2} > 0$ since, by assumption, $\tau > 0$. Here, c_1, c_2 are constants determined by the initial data. Hence, if $c_2 \neq 0$, the solution grows exponentially in the future.

On the other hand one may obtain solutions to (11) as follows. Substituting, $y = e^{zt}$ (z complex) in (11), and cancelling e^{zt} from both sides, we find that z must satisfy

$$z = e^{-z\tau} - 1, \quad \text{or} \quad z + 1 = e^{-z\tau}. \quad (15)$$

This is the so-called characteristic quasi-polynomial equation. For large $|z|$ we can approximately replace $z + 1$ by z, so this reduces approximately to

$$z = e^{-z\tau}, \quad \text{or} \quad ze^{z\tau} = 1. \quad (16)$$

The above quasi-polynomial equation admits an infinity of roots. Without going into the detailed derivation, these roots (hence also the large modulus roots of (15)) are approximately given by[3]

$$z_k \approx -\frac{1}{\tau}\ln(2k - \frac{1}{2})\frac{\pi}{\tau} + i(2k - \frac{1}{2})\frac{\pi}{\tau}, \quad (17)$$

where k is an integer. Since $\tau > 0$ has been assumed small, the above roots z_k of the quasi-polynomial all lie in a left half-plane $\text{Re}(z_k) < 0$. Each root z_k corresponds to a solution of (11) of the form $y = e^{z_k t}$, or to an oscillation with amplitude $e^{\text{Re}(z_k)}$. If $\text{Re}(z_k) < 0$, this solution must be exponentially damped. Hence, there are an infinity of solutions of (11) which are exponentially damped oscillations, contrary to the exponentially increasing solutions (14) of the approximating ODE (13).

Thus, expanding FDEs to obtain ODEs, by means of a "Taylor's" series, may lead, as above, to spurious solutions with the completely opposite behaviour. Note, also, in passing, that the retarded FDE (11) admits an infinity of distinct complex solutions which is impossible for the approximating ODE (14), or any ODE.

Electrodynamics provides a common situation where FDEs arise in physics, and where such a mistaken conversion of FDEs to ODEs is common.

3 The motion of two charges

3.1 Preliminaries

Consider *two* interacting charges which are otherwise isolated. According to present-day physics the cosmos consists mostly of charged particles, such as the electron, proton, etc. So the interaction of two charges should be simple problem which every physicist should know how to solve!

However, physics students are taught in high school that this basic problem requires quantum mechanics; they are taught that it *cannot* be solved using classical electrodynamics. Exactly why not?

What I will now show is this: the belief that classical electrodynamics is inadequate may be right or wrong, but it is based on bad reasoning. (And even if we arrive at the right answer for wrong reasons that is not science.) Classical electrodynamics actually leads to FDEs which physicists mistakenly converted to ODEs to draw the conclusion that classical electrodynamics does not work for the atom.

The FDEs for the classical electrodynamic 2-body problem are explicitly written down in [4](relativistic case) and [5] (non-relativistic case). Here let us intuitively understand why at all FDEs are needed. On classical electrodynamics, the two charges interact through electromagnetic fields. Each charge moves in the electromagnetic field of the other.

3.2 Heaviside-Lorentz force

A charge moving in an electromagnetic field experiences a Heaviside-Lorentz force given by the expression:

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}). \tag{18}$$

From this force, we can determine the motion of the charge by applying Newton's law of motion. These two laws together determine the motion of each charge if the electromagnetic field of the other is known.

3.3 Maxwell's equations

The electromagnetic field of each charge is determined using Maxwell's equations:

$$\begin{aligned}
\nabla \cdot \vec{E} &= \frac{\rho}{\epsilon_0}, \\
\nabla \cdot \vec{B} &= 0, \\
\nabla \times \vec{E} &= -\frac{\partial B}{\partial t}, \\
\nabla \times \vec{B} &= \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial E}{\partial t}.
\end{aligned} \tag{19}$$

We must solve these PDEs for each charge.

How to do so is explained in any standard physics text[6]. The fields $\vec{E}$ and $\vec{B}$ are calculated as the derivatives of a scalar potential V and a vector potential $\vec{A}$.

$$\begin{aligned}
\vec{E} &= -\nabla V - \frac{\partial \vec{A}}{\partial t} \\
\vec{B} &= \nabla \times \vec{A}
\end{aligned} \tag{20}$$

With this, the middle two of Maxwell's equations are automatically satisfied. The choice of potential is not unique, and we can add an extra condition (called a gauge condition) to simplify the remaining two equations. In the Lorenz gauge, the first and the last of Maxwell's equations turn into inhomogeneous wave equations for the scalar and vector potential respectively.

3.4 The Lienard-Wiechert potentials

We can solve the inhomogeneous wave equation for any charge distribution, provided we know the solutions for a point charge, or a δ

function charge density. (Such solutions are also known as Green functions, and mathematicians call them fundamental solutions.) We can calculate these Green functions by taking a Fourier transform (the Fourier transform of the delta function or distribution is 1), solving the resulting algebraic equation, and then applying the inverse Fourier transform.

The potentials V and $\vec{A}$ in this case are known as the Lienard-Wiechert (L-W) potentials. Unlike the Newtonian gravitational potential, which propagates instantaneously, these L-W potentials propagate only with the speed of light c. Thus, the field acting on charged particle A *now* $(t = 0)$ depends upon the motion of B at a *different* time $(t \neq 0)$.

Which different time? That depends on which kind of the L-W potential we use, for the L-W potentials are of two kinds—retarded and advanced. In the retarded case, the field acting on a charge A now depends upon the motion of B in the *past* $(t \leq 0)$. In the advanced case it depends upon the motion of B in the *future* $(t \geq 0)$. Usually, only the retarded potentials are considered on grounds of "causality". Let us go along with that for the moment.

The retarded L-W potentials are given by the expressions:

$$V(\vec{r},\, t) = \frac{1}{4\pi\epsilon_0}\, \frac{qc}{(Rc - \vec{R}\cdot\vec{v})}\Bigg|_{\text{ret}}, \qquad \vec{A}(\vec{r},\, t) = \frac{\vec{v}}{c^2}\, V(\vec{r},\, t)\Big|_{\text{ret}}. \tag{21}$$

This give us the potentials (hence the fields) at any position $\vec{r}$ at any time t, due to a charge q. Here, $\vec{R} = \vec{r} - \vec{r}_p(t_r)$, where $\vec{r}_p(t_r)$ denotes the position of the charge q at the *retarded time* t_r. The subscript "ret" emphasizes that $\vec{v}$, the velocity of the charge q, is also to be evaluated at retarded time t_r, so that $\vec{v} = \dot{\vec{r}}_p(t_r)$. The retarded time t_r satisfies

$$c^2(t - t_r)^2 = R^2, \tag{22}$$

where c is the speed of light.

3.5 Retarded time

Exactly what is this retarded time? The equation (22) is actually the equation of the null cone. Suppose we look at particle B from the position of A at time $t = 0$. The light wave from B which reaches us just now $(t = 0)$ started off in the past, from the "last seen" position of B. Light waves travel along the null cone, so the last-seen position of B is its retarded position. The corresponding time is the retarded time. The position, velocity etc. of B at *retarded time* is what we must use to evaluate the L-W potentials and calculate the fields acting on A now (and vice versa).

Geometrically, the retarded time is obtained as follows. We construct the backward null cone with vertex at the worldline of A at $t = 0$, and find the point, $t = t_B$ say, at which it intersects the worldline of B. That is equivalent to saying that the forward null cone with vertex at $t = t_B$ on the worldline of B intersects the worldline of A at $t = 0$. Although the point t_B is only in the *relative* past of $t = 0$, recall that the solution of the FDEs needs past data, or knowledge of the past worldline. Hence knowledge of the absolute past is needed to solve the resulting equations.

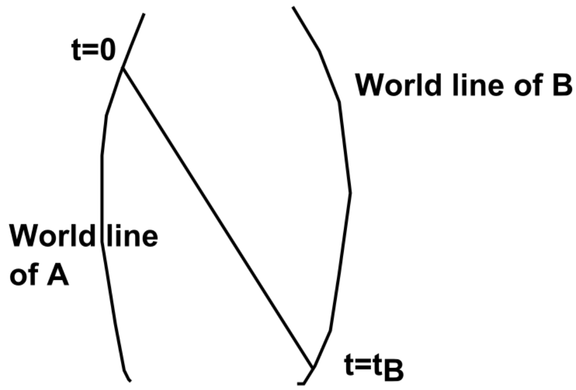


Figure 4: **R**etarded time. The retarded time is obtained by drawing the backward null cone from the present position of A and finding the point at which it intersects the world line of B.

This means that for a system of two charged particles, the future motion is not decided by their states at the present moment alone—we need to know their past motions. Since, charged particles (electrons, protons, ...) are everywhere that means we have a paradigm shift in physics: according to *existing* physics. We have *not* introduced any new hypothesis, but have just done the math correctly. The conclusion is that, contrary to common belief, the future states (of a classical system) are *not* decided by its present state alone; we require knowledge of its past.

4 The Groningen debate

This conclusion about a paradigm shift in physics was first published in this very journal in 1992, as part of a series of 10 articles "On Time" which were published between 1990 and 1994. The intention was to carry on the conversation which used to take place under the tree near the Pune university canteen.

However, many scientists today assess the validity of a scientific claim, not by applying their mind to the claim, but just by judging the "prestige" of the publisher who published it. Though this regrettable practice of judging scientific truth by social prestige and authority is contrary to the spirit of science, it is widely prevalent. Accordingly, I republished this conclusion about a paradigm shift (and the whole related series of papers) as a book, *Time: Towards a Consistent Theory*, with a "prestigious" foreign publisher (Kluwer, now Springer), in 1994,[7] since every scientist knows that only Western publications are prestigious!

Nevertheless, this claim about a paradigm shift ran into heavy fire, when I mentioned it at a meeting ("Retrocausality Day") in Groningen, in 1999. One of the participants, H. D. Zeh, a professor from Heidelberg, objected strenuously, asserting that no paradigm shift was needed. The whole meeting was stalled by the debate on this one point: is there or is there not a paradigm shift?

Zeh's argument was this: the motion of two charges is completely decided by two equations.

1. The Heaviside-Lorentz force (18), which force gives us a system of ODEs for the motion of each charge, by Newton's second law of motion.

2. Maxwell's equations (19) which deter-

mine the fields E and B generated by each particle (to be plugged into (18)). These are PDEs.

Hence, Zeh argued, to solve for the motion of the two charges we only need to solve ODEs and PDEs, never any FDEs. Further, Zeh argued, for either ODEs or PDEs initial data is adequate. Therefore, Zeh concluded, there is no need for past data, and consequently no paradigm shift.

My immediate response was that I had actually solved the electrodynamic 2-body problem using FDEs. This solution required past data, and if Zeh had a different way of obtaining the solution, he ought to show it. Zeh replied that he was sure of his stand on "physical grounds", and that actually solving the equations was the mathematician's job!

The puzzle which emerged was this: my formulation of the electrodynamic 2-body problem in terms of FDEs ultimately used nothing more than the very same Heaviside-Lorentz force (ODEs) and Maxwell's equations (PDEs). So how could two such very different conclusions emerge from the *same* underlying physics? Is there, or is there not a paradigm shift involved? Is there, or is there not a need for past data in physics?

Though the participants in the meeting came from many prestigious universities around the world, no one had a ready answer. The debate remained unresolved during the Groningen meeting. We will see the resolution in the next part.

References

[1] R. D. Driver. *Introduction to Differential and Delay Equations.* Springer, Berlin, 1977. The existence theorem there does not, however, apply to FDEs with unbounded delays, which *may* arise in electrodynamics.

[2] C. K. Raju. Thermodynamic time. In *Time: Towards a Consistent Theory*, chapter 4, pages 79–101. Kluwer Academic, Dordrecht, 1994.

[3] L. E. El'sgol'ts and S. B. Norkin. *Introduction to the theory of differential equations with deviating arguments.* Holden-Day, San Francisco, 1966. Typo corrected in equations (28) and (29) on p. 74.

[4] C. K. Raju. Electromagnetic time. *Physics Education (India)*, 9(2):119–128, 1992. http://arxiv.org/pdf/0808.0767v1.

[5] C. K. Raju. The electrodynamic 2-body problem and the origin of quantum mechanics. *Foundations of Physics*, 34(6):937–962, 2004. http://arxiv.org/pdf/quant-ph/0511235v1.

[6] David J. Griffith. *Introduction to Electrodynamics.* Prentice Hall, India, 1999.

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

Functional differential equations. 2: The classical hydrogen atom

C. K. Raju

AlBukhary International University
05200 Alor Setar, Malaysia
ckr@ckraju.net

(Submitted 22-07-2013)

Abstract

The previous part closed with a doubt: textbook electrodynamics mentions only ordinary and partial differential equations (ODEs and PDEs), never any functional differential equations (FDEs). So are FDEs or past data really needed? In fact, to solve Maxwell's PDEs we need Cauchy data on fields, which is equivalent to past data on particle motions. Electrodynamics actually involves a *coupled* system of ODEs and PDEs, equivalent to FDEs which are currently the best way to solve such a coupled system. The full electrodynamic force is commonly approximated by the Coulomb force, but this approximates FDEs by ODEs, and is error prone. Specifically, circular or central orbits are not valid solutions of the electrodynamic 2-body problem, in the *absence* of radiation damping. We can see this heuristically, for the classical hydrogen atom, where the full electrodynamic force involves a delay torque. Hence, the century-old argument that classical electrodynamics is inadequate is based on erroneous reasoning (irrespective of whether the conclusion is valid).

1 Recap

To recapitulate the first part, the motion of two charges is determined by the Heaviside-Lorentz force and Maxwell's equations. This leads to retarded FDEs. Such FDEs present a paradigm shift in physics, because past data is needed to solve them. However, the need for such a paradigm shift (or the need to prescribe past data) was contested in the Groningen debate. Zeh argued that the Heaviside-Lorentz force (and Newton's second law) lead to ODEs, while Maxwell's equations are PDEs. For solving either ODEs or PDEs, only initial data is needed. Hence, he maintained, there is no need for past data, and no paradigm shift. The puzzle was that the same underlying physics (Heaviside-Lorentz force, and Maxwell's equations) seems to lead to two opposite conclusions about the need for past data.

2 The resolution of the Groningen debate

Subsequently, I gave a simple resolution of this puzzle, which was eventually published in the *Foundations of Physics* 2004[1] which had Zeh on its editorial board. The simple resolution which relates FDEs to ODEs and PDEs is as follows.

2.1 Hyperbolic PDEs

For ODEs initial data determines a unique solution, as we know, and Peano's existence theorem provides a formal proof. For PDEs the situation *seems* the same, but is actually a little more complicated.

First, initial data are appropriate only for certain *types* of PDEs, called hyperbolic PDEs. The wave equation is a canonical example of a hyperbolic PDE: the effect (the wave) depends upon a specific cause (the source) in the past. The wave travels out and gets disconnected from the source. In contrast, the potential equation (Laplace equation, Poisson equation) is an example of an elliptic PDE. An elliptic PDE involves action at a distance: the Newtonian gravitational potential or the Coulomb field is decided everywhere and for all time by specifying the position of its source. The field remains attached to its source, so that boundary data are more appropriate than initial data.

2.2 Cauchy data

Initial data—formally called Cauchy data—involve values of the function (and its partial derivatives for a higher order equation). These must be specified on an appropriate hypersurface, called a Cauchy hypersurface. [A hypersurface is an $n-1$ dimensional subspace of an n dimensional space. Thus, a plane is a hypersurface in 3-dimensions, while 3-d space is a hypersurface in 4-dimensional spacetime.] Prescribing Cauchy data for a PDE is *analogous* to prescribing initial data for an ODE, since the "Cauchy hypersurface" may be just the "present instant", or the hypersurface in spacetime given by the equation $t = 0$. However, the analogy may be misleading.

2.3 What Cauchy data is needed for Maxwell's equations?

Many PDEs of physics (like the Hilbert-Einstein equations or Navier-Stokes equations) are not strictly hyperbolic, but let us just assume that Maxwell's equations are hyperbolic for (in the Lorenz gauge) they are just inhomogeneous wave equations.

Exactly *what* Cauchy data do we need for

Maxwell's equations? Since those equations involve the derivatives of the fields E and B, we must prescribe E and B on the Cauchy hypersurface, say $t = 0$. **That is, to solve Maxwell's equations, we must prescribe the electric and magnetic fields over *all* space at one instant of time.** But how do we *do* that?

2.4 Cauchy data for fields = past data for particles

In our case of two charged particles, consider a point P on the Cauchy hypersurface $t = 0$. The field at P must come from the motion of the two particles, for they are the only sources of fields by assumption. Since we have assumed retarded potentials, the field must come from their *past* motions. Which past motions? To determine the field at P, we must again construct the backward null cone with vertex at P and determine the two points (say P_A, and P_B) where it intersects the worldlines of the two particles, A, and B.

As the point P moves further away in space, the corresponding points P_A and P_B (at which the backward null cone from P intersects the two world lines) will move further back in time. As P runs over the entire hypersurface, the corresponding points P_A and P_B will cover the entire past world lines of the two particles. That is, (assuming retarded propagators) prescribing Cauchy data for fields on a hypersurface (i.e., at one instant of time) is equivalent to prescribing *past* data on the particle world-lines.

That is to solve Maxwell's equations, we need to prescribe the electric and magnetic fields over all space at only *one* instant of time. *But to do that we need to know the past motions of the two particles for all time.*

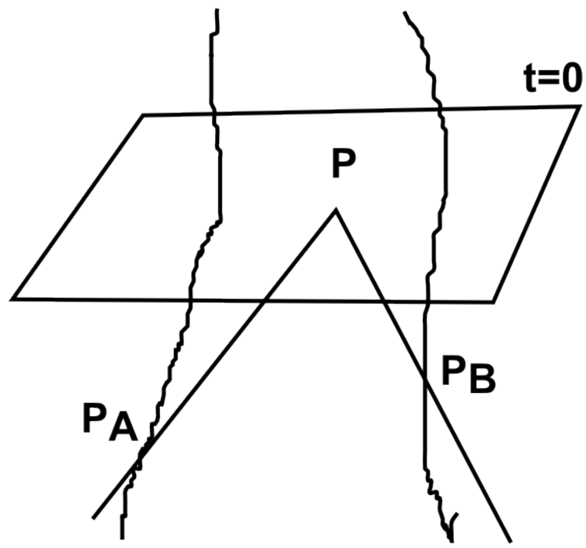


Figure 1: **I**nitial data for fields is past data for particles. To solve Maxwell's equations we must prescribe the fields $\vec{E}$ and $\vec{B}$ over all space at one instant of time $t = 0$. For the 2-body problem the field at a point P on $t = 0$ depends on the past motions of particles A and B at the retarded positions P_A and P_B.

So, past data *is* needed to solve for the motion of two charges, whichever way we look at it: whether in terms of fields or particles. Because the mathematical theory of PDEs is a little more complicated than that of ODEs, and many physicists are not very familiar with it, they just unthinkingly extended the Newtonian paradigm of ODEs to PDEs.

Further, physicists have unfortunately got carried away by the intuitive picture of a field.

The particle pictures pins us down and reveals the reality of retarded action at a distance (implicit in the concept of hyperbolicity). "Initial data" for fields (or knowledge of the fields now) involves past data on particle motions (or knowledge of the entire history of the two particles).

That is, Zeh made two mistakes: he put his faith in the Newtonian paradigm ("initial data are enough"), and he neglected the details of the actual calculations involved ("that is the mathematician's job"), which would have made manifest the need for past data. One should clearly understand that this is not about Zeh as a person: he only personifies and represents a mistake made by practically the entire physics community, a mistake which has persisted for over a century.

2.5 Coupled ODEs and PDEs

One further point needs to be noted. The ODEs of motion (Heaviside-Lorentz force law + Newton's second law) determine the motion of the charges *if the fields are known*. Maxwell's equations (PDEs) determine the fields *if the motion of the charges are known*. So, to solve for the motion of two charges, we have to solve both *together*. So, what we have here is a *coupled* system of ODEs and PDEs, not an independent pair of ODEs and PDEs.

Unlike Peano's existence theorem for ODEs, or the Cauchy-Kowaleski theorem for PDEs, there is still no corresponding formal mathematical theorem for a *coupled* system of ODEs and PDEs, as of now. (Though it is easy to prove such a theorem, I won't do it, since, I now work with a different philosophy of mathematics,[2] and no longer think formal mathematical proofs are important.)

I mention this only to bring out that the coupled system of ODEs and PDEs is distinct from the two systems of ODEs and PDEs considered individually. It is necessary to hammer home the point about the coupled system, since most current texts on electromagnetic theory do not consider this coupling at all.

It is this coupled system of ODEs and PDEs which is equivalent to FDEs. Briefly:

$$\text{ODEs} + \text{PDEs} = \text{FDEs}.$$

We can solve either the coupled system of ODEs and PDEs or we can solve FDEs. However, FDEs are much easier to solve numerically, and well-tested computer programs to do so have been readily available for a quarter century.[3] (Also, as stated earlier, there are formally proven theorems that solutions of FDEs exist with past data.) Whichever way we look at it, it is, at present, preferable to solve FDEs instead of a coupled system of ODEs and PDEs, though the two are equivalent.

Either way, the paradigm shift is there, it is real, and it comes not from any new physical hypothesis but from a better understanding of the mathematics—an understanding which has been sadly missing for a whole century.

3 The classical hydrogen atom

So, what actual difference does that make to physics? Let us again consider our two charges, but this time in the explicit context of the classical hydrogen atom. Almost exactly a century ago, it was declared that classical electrodynamics does not work for this case, and this argument would be familiar to most physics students from high school.

3.1 The text-book story

Under the weight of the Newtonian paradigm, the atom was regarded as a sort of a miniature solar system, for the inverse-square law Coulomb force is just like the inverse-square law Newtonian gravitational force. Usually a further simplification is made: the proton is regarded, like the sun, as infinitely massive, hence unmoving at the centre. This reduces the 2-body problem to a *1-body case* of central orbits. Physics students are very familiar with central orbits which they study as part of their undergraduate course. According to the theory of central orbits, circular (or elliptic orbits) are stable with an inverse square law central force.[4]

Beyond this point, students learn, the analogy between classical electrodynamics and Newtonian gravitation fails: an accelerating charged particle gives out electromagnetic radiation. And, motion in a central orbit involves constant acceleration. Thus, it is argued, the classical central orbits are not stable in the electrodynamic case but must decay because of the associated radiation damping. The electron would constantly lose energy and eventually spiral into the nucleus. Hence the conclusion that classical electrodynamics cannot correctly describe the hydrogen atom.

3.2 The full electrodynamic force

It is odd how physicists have accepted this argument for over a century, when it is immediately obvious that something is wrong with it. The Coulomb potential propagates instantaneously like the Newtonian gravitational potential, but the actual Lienard-Wiechert (L-W) electromagnetic potentials propagate only at the speed of light. To correct this, we must use the full electrodynamic force, not just the Coulomb force.

The full electrodynamic force is obtained as follows. Recall that the electric and magnetic fields are obtained by differentiating the scalar potential V and the vector potential $\vec{A}$:

$$\begin{aligned} \vec{E} &= -\nabla V - \frac{\partial \vec{A}}{\partial t}, \\ \vec{B} &= \nabla \times \vec{A}. \end{aligned} \tag{1}$$

We need to apply this to the L-W potentials

$$\begin{aligned} V(\vec{r},\ t) &= \frac{1}{4\pi\epsilon_0}\, \frac{qc}{(Rc - \vec{R}\cdot\vec{v})}\bigg|_{\text{ret}}, \\ \vec{A}(\vec{r},\ t) &= \frac{\vec{v}}{c^2}\, V(\vec{r},\ t)\bigg|_{\text{ret}}. \end{aligned} \tag{2}$$

Recall that $\vec{R} = \vec{r} - \vec{r}_p(t_r)$ is the vector connecting the spacetime point $(\vec{r},\ t)$ to the position $\vec{r}_p(t_r)$ of charge q at retarded time t_r, which satisfies $c^2(t - t_r)^2 = R^2$. Recall also that the velocity too must be evaluated at retarded time, $\vec{v} = \dot{\vec{r}}_p(t_r)$.

Carrying out the differentiation is a long and tedious process, found in standard texts.[5] This gives us the following expressions for the electric and magnetic fields.

$$\vec{B}(\vec{r},t) = \frac{1}{c}\hat{\vec{R}} \times \vec{E}(\vec{r},t),$$
$$\vec{E}(\vec{r},t) = \frac{q}{4\pi\epsilon_0}\frac{R}{(\vec{R}\cdot\vec{u})^3}\Big[(c^2-v^2)\vec{u} + \vec{R}\times(\vec{u}\times\vec{a})\Big]. \quad (3)$$

Here, we have introduced $\vec{u} = c\hat{\vec{R}} - \vec{v}$, and $\vec{a}$ is the acceleration at retarded time: $\vec{a} = \ddot{\vec{r}}_p(t_r)$.

If the (retarded) acceleration is zero, we are left with only the first (velocity dependent) term for $\vec{E}$. If the retarded velocity too is zero, this term reduces to a term similar to the Coulomb force:

$$\vec{E}(\vec{r},t) = \frac{1}{4\pi\epsilon_0}\frac{q}{R^2}\hat{\vec{R}}. \quad (4)$$

(It is exactly the Coulomb force if the charge q is static for all time.)

From the above expressions (3) for $\vec{E}$ and $\vec{B}$, we can calculate the force on a charge q_1 moving with velocity $\vec{v}_1$ using the Heaviside-Lorentz force law: $\vec{F} = q_1(\vec{E} + \vec{v}_1 \times \vec{B})$. This gives the expression:

$$\vec{F} = \frac{qq_1}{4\pi\epsilon_0}\frac{R}{(\vec{R}\cdot\vec{u})^3}\Big\{\Big[(c^2-v^2)\vec{u} + \vec{R}\times(\vec{u}\times\vec{a})\Big] + \frac{\vec{v}_1}{c}\times\Big[\hat{\vec{R}}\times[(c^2-v^2)\vec{u} + \vec{R}\times(\vec{u}\times\vec{a})]\Big]\Big\}. \quad (5)$$

3.3 Full electrodynamic force leads to FDEs

Using the expression (5) for the full electrodynamic force leads to FDEs because of the retarded quantities involved. If we use only the Coulomb force, that would lead to ODEs. So, approximating the full electrodynamic force by the Coulomb force amounts to approximating FDEs by ODEs. That is an error-prone process, as already explained in the first part of this article.

On the other hand, if we did not solve those FDEs, how do we know that circular orbits are stable? Can we just say, like Zeh, that we know this "on physical grounds" and that it is the mathematician's job to prove us right?

It is not very difficult to numerically solve the retarded FDEs for the non-relativistic case without radiation damping. Nevertheless, for the classical hydrogen atom, the first solution was published by me only in 2004 while resolving the Groningen debate. Here are some of the problems that arise.

3.4 How should past data be prescribed?

The first problem is the question of prescribing past data. How exactly should that be done? Because the issue of past data has not been considered in physics until now, there are no clear guidelines available. However, the above connection between FDEs and PDEs suggests a way: prescribing past data on particle motions is equivalent to prescribing fields on a spacelike Cauchy hypersurface. Allowing arbitrary Cauchy data corresponds to allowing arbitrary past motions of the two particles. That is, we don't try to explain how those particle motions came about.

Fortunately, unlike the PDE case, for which one must (on existing theory) prescribe the *entire* past history, for the FDE case it is usually enough to prescribe only a short portion of the past history of the two particles, since we are typically interested in the solution for a short period of space and time. The same can be done with ODEs+PDEs, but there is no formal proof for that, as of now.

However, various questions arise. How did that past history come about? Is it physically realizable? These are questions which the mathematical theory of FDEs and PDEs cannot answer, and which physicists have not raised or answered till now. Prescribing past data arbitrarily may lead to a discontinuity in the solution or (more usually) its derivatives. This is not a major technical problem, since most existing numerical codes are equipped to handle such discontinuities. The discontinuity may be physically understood as arising from suddenly "switching on" the interaction. Nevertheless, this is an unsatisfactory state of affairs. Can the past data be prescribed in such a way as to avoid this discontinuity? The answer to this question is not known at present.

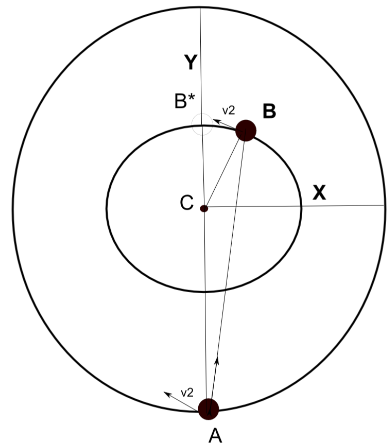


Figure 2: **E**ffect of retardation in the 2-body problem. Classically, both particles A and B rotate around a common centre of mass, C. The force acting on A now involves the past or retarded position of B and not its instantaneous position ($B*$). This force does not pass through the instantaneous centre of mass and hence results in a torque, called the "delay torque".

4 The delay torque

Anyway, without going into all the details of that solution, we can heuristically see the problems involved.

Examine Fig. 2. First, note the difference from central orbits (or the 1-body problem). For the *two* body problem, even in classical Newtonian gravity, both bodies revolve around a common centre of mass. To put matters in another way the motion of planets is *not* heliocentric, but is better understood as barycentric.

The next step is to take into account the retardation, since the full electrodynamic force does not travel instantaneously but only at the speed of light.

Consider now just the first term in the first square bracket in (5) which corresponds to the "Coulomb" component of the electric field due to q. This component acts in the direction of the vector $\vec{u}$, given by $\vec{u} = c\hat{\vec{R}} - \vec{v}$. If $\frac{v}{c}$ is small this is approximately the same direction as $\hat{\vec{R}}$. That is the force on A (charge q_1) due to B (charge q) will point approximately towards the "last seen" position of B, not its instantaneous position. Consequently, the force will not pass through the instantaneous centre of mass (barycentre) of the two particles and it will exert a torque on particle A. That means that circular orbits will *not* be stable even in the *absence* of radiation damping!

Of course, we should take into account that the force acts in the direction of the vector $\vec{u}$ and not just the vector $\vec{R}$. But will that resolve the problem in an obvious way? Will that "correct" the force, so that it points towards the instantaneous centre of mass? It cannot because there is no physical way to ascertain the exact instantaneous position of B, hence no way to ascertain the instantaneous centre of mass: at best the full force may point towards the estimated or extrapolated instantaneous centre of mass. That is of little consequence, for that process of extrapolation would fail for more complex motions. We will consider this argument again, in the context of gravity.

4.1 Choice of gauge

Have we neglected anything else? What if we were to use the Coulomb gauge instead of the Lorenz gauge? Choices of gauge typically confuse physics students, for the scalar potential in the Coulomb gauge is just the same old Coulomb potential which propagates instantaneously. However, that is deceptive: changing the gauge does *not* change the electric or magnetic fields. It does not change the full force in (5). So the torque, too, does not depend upon the gauge.

4.2 Delay torque initially accelerates the particle

A glance at Fig. 2 shows that the torque will *accelerate* the electron. We can confirm this by evaluating the delay torque numerically for an electron initially in a classically stable orbit: the torque (initially) accelerates the electron. That is, *in the absence of radiation damping*, an electron initially in a classically stable orbit (on the Newtonian paradigm) tends to fall *out* of the atom!

The exact motion may be complicated, but the point here is only this. The whole argument given by Bohr for the instability of the classical hydrogen atom, an argument repeated for a century in physics texts, is defective because it assumed (on the Newtonian paradigm) that classical central orbits are stable in the absence of radiation damping. (To reiterate, it does not matter whether the conclusion is still valid, for even if one gets the right answers for the wrong reasons, that is not science.)

Further, we need to take into account motions more complex than simple circular orbits, for it may be that a more complex past motion (such as an oscillation superposed on a circular orbit) leads to stable solutions. Whether or not that is so, is an open question at present. We cannot decide that without solving the FDEs. Unfortunately, this has not been done for the past century.

5 Summary

Thus, for the 2-body problem of classical electrodynamics, the stability of orbits, even in the absence of radiation damping, is a complex problem which has not been properly studied so far. The whole physics community just went along with a wrong solution based on the Newtonian paradigm.

A natural question arises. If, in the absence of radiation damping, the delay torque makes circular orbits unstable, so that the electron tends to fall *out* of the atom, then will stability be somehow restored by reintroducing radiation damping? In short, are there motions for which the delay torque and the radiation damping cancel (either exactly or on an average)? We will examine this question in the next part which connects radiation damping to FDEs.

References

[1] C. K. Raju. The electrodynamic 2-body problem and the origin of quantum mechanics. *Foundations of Physics*, 34(6):937–962, 2004. http://arxiv.org/pdf/quant-ph/0511235v1.

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

[3] E. Hairer, S. P. Norsett, and G. Wanner. *Solving Ordinary Differential Equations*. Springer, revised ed. 1991 edition, 1987.

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

[5] David J. Griffith. *Introduction to Electrodynamics*. Prentice Hall, India, 1999.

Simulation of three-level lasing system using Cuboctahedron

Madhuramurthy M N [1] and Sarmistha Sahu [2]

[1] II BSc student, [2] Head, Department of Physics,

Maharani Lakshmi Ammanni College for Women,

Science Post, Bangalore 560012 (INDIA)

sarmistha.sahu@gmail.com

(*Submitted* 19-09-2013)

Abstract

Einstein's constants and its importance in understanding the principles of laser technology are best understood in this play-learn mode. Dice throwing is a pedagogical tool for learning probabilities. The aim of the experiment is

- to determine Einstein's constants for a lasing system.
- to find life time of metastable state and excited state of an atom.
- to show why 3-level lasing system is preferred to two level lasing systems.

Keywords: Einstein's constants, Population inversion, optical pumping, lifetime, two level lasing system, dice throwing, cuboctahedron, three level lasing system, metastable state.

1. Introduction

Dice rolling is being used as a pedagogical tool in schools as well as undergraduate studies in physics, mathematics, statistics and computer science curriculum. Students learn while they play. Todd W. Neller et al [1] has used dice rolling in a dice game Pig for undergraduate research in machine learning. Arthur Murray et al [2] mention that "the 'radioactive dice' experiment is a commonly used classroom analogue to model the decay of radioactive nuclei". Simple dice rolling can unfold important concepts elegantly, in fact, this experiment explains all the keywords mentioned above.

2. Dimensional Analysis

The Einstein constant A is defined as the *probability per unit time*. And the product of the Einstein constant B and the energy density per unit volume u of the stimulant is the *probability per unit time* (refer to Eqs. (3) and (1) respectively)

Rolling the dice and watching for a particular "face up" gives the *probability per throw*, and therefore can represent the Einstein's constants A and Bu.

Thus, simulation of dice with atoms, time with throw and three faces of the dice with the three processes help in understanding the lasing operation.

3.Theory

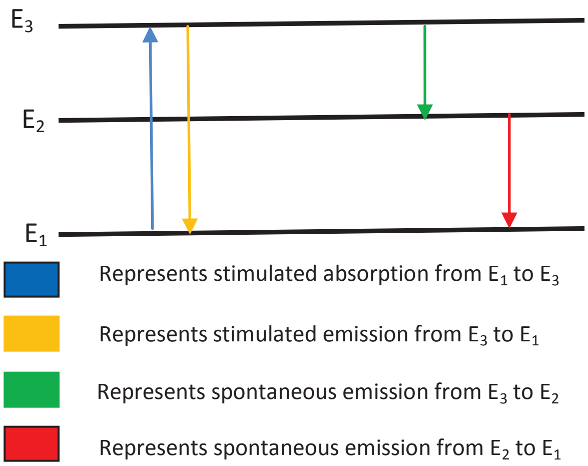


Fig **1. Three level system**

For a three level system with population N_1, N_2 and N_3 in the energy levels E_1, E_2 and E_3 respectively and a stimulant of energy density $u(\omega)$ where $\omega = \frac{E_3 - E_1}{\hbar}$,

the number of stimulated absorptions per unit volume per unit time from level 1 to 3 is

$$\frac{dN_3}{dt} = -\frac{dN_1}{dt} = N_1 B_{13} u(\omega) \qquad (1)$$

The number of stimulated emissions per unit volume per unit time from level 3 to 1 is

$$\frac{dN_3}{dt} = -N_3 B_{31} u(\omega) \qquad (2)$$

The rate of spontaneous transitions (per unit volume) per unit time from level 3 to 2 is proportional to N_3 and is given by

$$\frac{dN_3}{dt} = A_{32} N_3 = -\frac{N_3}{t_{sp3}} \quad \ldots\ldots\ldots\ldots(3)$$

The rate of spontaneous transitions (per unit volume) per unit time from level 2 to 1 is proportional to N_2 and given by

$$\frac{dN_2}{dt} = -A_{21} N_2 = -\frac{N_2}{t_{sp2}} \qquad (4)$$

The **life time** of the level 2 is

$$t_{sp2} = \frac{1}{A_{21}}$$

and that of th

e level 3 is

$$t_{sp3} = \frac{1}{A_{32}}$$

The quantities **B_{31}, B_{13}, A_{32}** and **A_{21}** are known as **Einstein's constants** and depends on the atomic system [3].

Thus, the rate equation for level 3 is,

$$\frac{dN_3}{dt} = -B_{13}(N_3 - N_1) - N_3 A_{32} \qquad (5)$$

the rate equation for level 2 is,

$$\frac{dN_2}{dt} = N_3 A_{32} - N_2 A_{21} \qquad (6)$$

the rate equation for level 1 is,

$$\frac{dN_1}{dt} = N_3 B_{31} + N_2 A_{21} - N_1 B_{13} \qquad (7)$$

For a system where $A_{21} < B_{13} = B_{31} < A_{32}$ (or $t_{sp3} < t_{sp2}$), at equilibrium,

$$\frac{dN_3}{dt} = 0 \;\; therefore \;\; N_3 = \frac{B_{13}N_1}{A_{32} + B_{31}} \approx \frac{B_{13}N_1}{A_{32}} \qquad (8)$$

$$\text{and} \;\; \frac{dN_2}{dt} = 0; \; therefore \; N_2 = \frac{A_{32}N_3}{A_{21}} \qquad (9)$$

Or using Eq(8) we get

$$N_2 \approx \frac{B_{31}N_1}{A_{21}} \qquad (10)$$

Since $B_{13} > A_{12}$, N_2 should be greater than N_1

4. Experimental procedure

4.1 Einstein constants

The cuboctahedron (dice) used is shown in Fig 2. Yellow face 'up' represents spontaneous emission from 2 $\rightarrow$ 1, Red face up for stimulated absorption/emission from $3 \leftrightarrow 1$ and the uncoloured face up for spontaneous emission from 3 $\rightarrow$ 2. The probability of spontaneous emission from the level 2 to 1 is obtained .by throwing (about) 100 dice and removing the ones with yellow-face-up

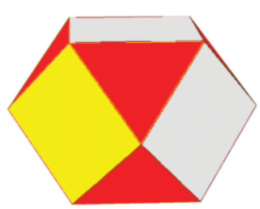
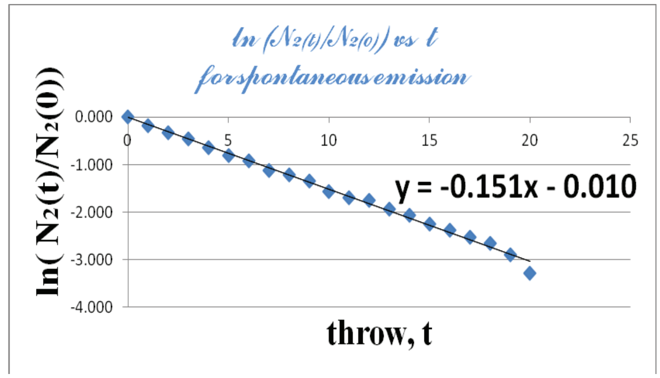


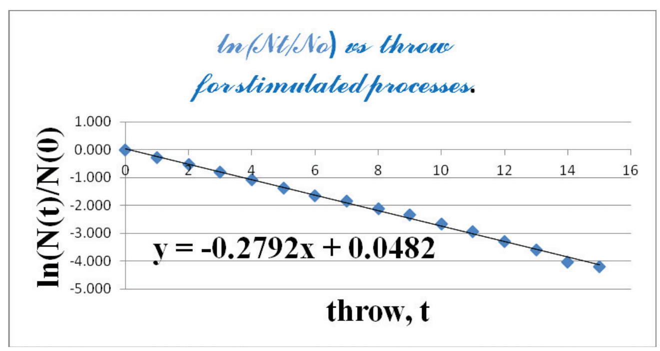
Fig 2. Cuboctahedron with three different faces: yellow, red and colourless.

(representing the atom that has undergone spontaneous emission from level 2 to 1). The number left $N_2(t)$ at each throw, t is played repeatedly till the die left is approximately 10% of the original. The plot of $\ln {}^{N_2(t)}/_{N_2(0)}$ *versus* t gives A_{21} as shown in Graph 1.[4,5]



Graph 1. Einstein constant A_{21} for spontaneous transition from 2 $\rightarrow$ 1 is (0.151±0.003) per throw and t_{sp2} =(6.67±0.013) throw.

Assuming that the Einstein's constants B_{13} is equal to B_{31}, the above procedure is repeated with the red-face-up to give B_{13} $u(\omega)$. B_{13} $u(\omega)$ is obtained from Graph 2.[4,5]



Graph 2: Einstein constant B_{13} multiplied with energy density $u(\omega)$ for the stimulated transitions $3 \leftrightarrow 1$ is (0.279 $\pm$ 0.014) per throw.

Therefore, the probability of the uncoloured-face-up representing A_{32} is given by Fig **2:**

$$A_{32} = 1 - A_{21} - B_{13}$$

$$= 1 - (0.151 \pm 0.003) - (0.279 \pm 0.014)$$

$$A_{32} = 0.570 \pm 0.017 \qquad (11)$$

4.2 Simulation of four processes in the 3-level lasing system with cuboctahedrons

In the cuboctahedron the probability of the yellow face up is 0.151, the red face up is 0.279 and the colorless face up is 0.570 as determined in the first part of the experiment and is used for representing A_{21}, $B_{13}u(\omega)$ and A_{32} respectively.

1. In the first throw, of the $N_3(0)$ dice (atom) from level 3, let $dN_{3st}(t)$ and $dN_{3sp}(t)$ be the red and white-faces-up respectively. They represent the stimulated and spontaneous emission from level $3 \rightarrow 1$ and $3 \rightarrow 2$ respectively. These dN_{3st} (t) and dN_{3sp} (t) dice are removed, to be left with $N_3(t)$ dice .
2. In the first throw, of the $N_2(0)$ dice in level 2, $dN_2(t)$ represents dice with yellow face up which have undergone spontaneous emission from level 2 $\rightarrow 1$. This $dN_2(t)$ dice are removed, to be left with $N_2(t)$ dice
3. Similarly, of the $N_1(0)$ dice of level 1; $dN_1(t)$ with the red-face-up, represents the *stimulated absorption* from level 1 $\rightarrow$ 3. These $dN_1(t)$ dice are removed, to be left with $N_1(t)$ dice .

In the (t+1) throw, the initial number of atoms (dice) in level 3 is given by

$$N_{3i}(t+1) = [N_3(t)+dN_1(t)] \qquad (12)$$

And the number of atom in the level 1 is

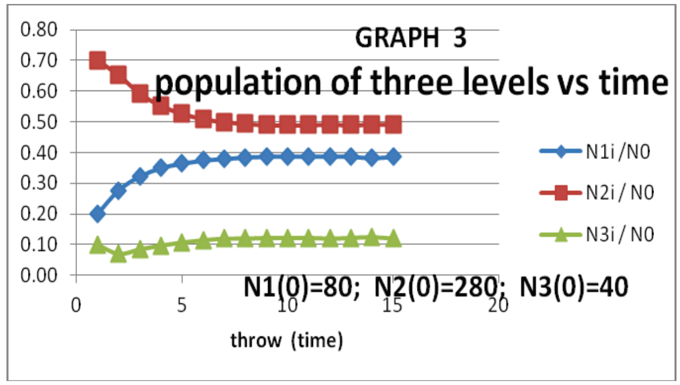
$$N_{1i}(t+1)=[N_1(t)+dN_2(t)+dN_{3st}(t)] \qquad (13)$$

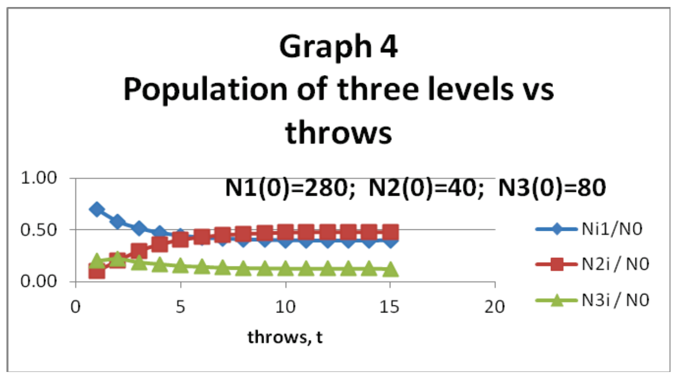
Similarly, in the (t+1) throw atoms at level 2 is

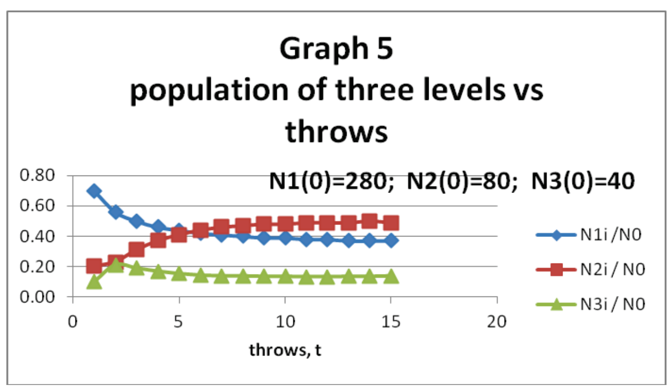
$$N_{2i}(t+1)=[N_2(t)+dN_{3sp}(t)] \qquad (14)$$

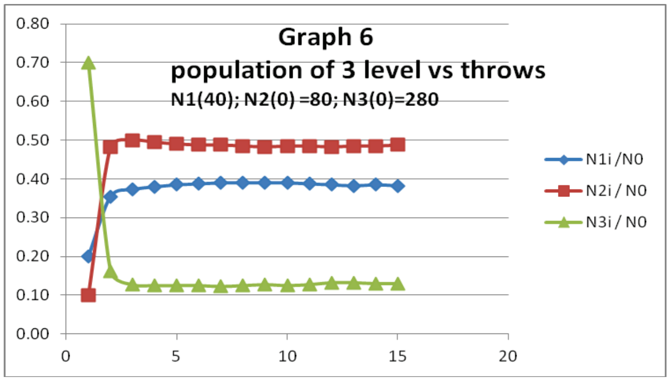
The above processes are repeated till the number of atoms in each level is stabilized.

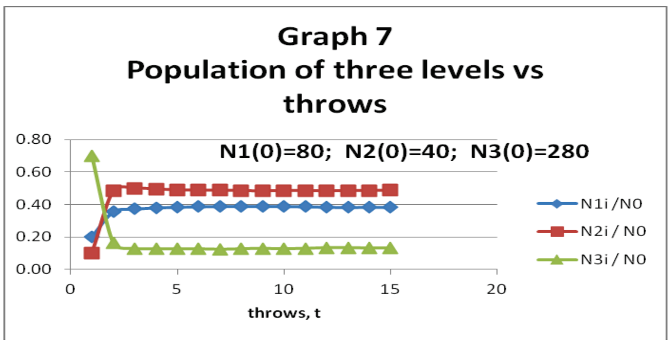
1. The plot of N_{1i}/N_o; N_{2i}/N_o; N_{3i}/N_o *vs t* is shown in Graphs 3, 4, 5, 6, 7, 8 for different initial populations.

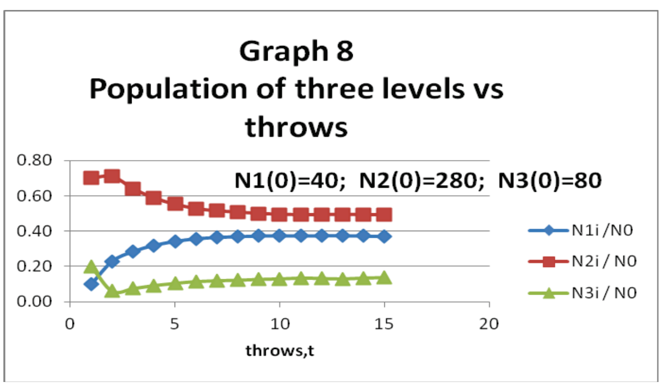












5. Results

1. From the graphs 3,4,5,6,7 and 8 we get

Graph no.	(N_{1i}/N_0)	(N_{2i}/N_0)	(N_{3i}/N_0)
3	0.39	0.49	0.12
4	0.40	0.48	0.12
5	0.38	0.49	0.13
6	0.39	0.48	0.13
7	0.39	0.48	0.13
8	0.39	0.48	0.13
Average—	**0.390**	**0.483**	**0.127**

2. From Graph 1, A_{21}= 0.151±0.003 per throw
3. From Graph 2, B_{13}=0.279±0.014 per throw
4. From Eq. (8)

$$N_3 = \frac{B_{13}N_1}{A_{32}+B_{13}}$$

Therefore,

$$A_{32} = \frac{B_{13}(N_1 - N_3)}{N_3}$$

$$= 0.578 \pm 0.013 \quad \text{...........(15)}$$

5. From Eq. (9)

$$N_2 = \frac{A_{32}N_3}{A_{21}}$$

Therefore,

$$A_{32} = \frac{N_2 A_{21}}{N_3} = 0.574 \pm 0.012 \quad \text{....(16)}$$

Thus, **A_{32} is 0.576±0.012 per throw.**

6. Inferences

1. The probability per unit throw of the spontaneous emission from level 2 to 1 is A_{21}=**0.151±0.003** per throw

2. The probability per unit throw of the spontaneous emission from level 3 to 2 is A_{32}=**0.576±0.012** per throw. This matches well with the assumed value of Eq. (12)

3 .The life-time of the atoms in the level 2 is t_{sp2} = (**6.67±0.013**) throw and level 3 is t_{sp3} = (**1.75±0.105**) throw. Hence $t_{3sp} < t_{2sp}$ .This represents level 2 of the system has a large life time as in a ***metastable*** state.

4. The product of the probability per unit throw of the stimulated absorption from 1 to 3 or stimulated emission from level 3 to 1 and the energy density is

$$B_{13}\, u(\omega) = B_{31}\, u(\omega) = \mathbf{(0.279 \pm 0.014)}$$

5 . The Population of the level 2 at equilibrium is always ***greater*** than that of level 1, when $N_2 > N_3 > N_1$, $N_2 > N_1 > N_3$, $N_3 > N_1 > N_2$, $N_3 > N_2 > N_1$, $N_1 > N_3 > N_2$, $N_1 > N_2 > N_3$ at the starting point.

Hence, amplification is possible and lasing action can take place.

6. A 3-level lasing system is preferred to two level since 2-level system[1] 2-level lasing systems are not good lasing material as there is no retention of population inversion for long, hence no amplification and therefore, no lasing action is possible

[1]Ms Madhura Murthy M N, Ms Reshma of II BSc and Sarmistha Sahu were awarded Prize in National Competition of Innovative Experiments -2012 for the experiment, '*Einstein's Constant A & B of 2-level lasing medium, Simulation using Cuboctahedron*' at the IAPT Convention at Cochin on 3 Nov 2012.
http://wikimediafoundation.org

7. Conclusion

Three or more levels, with at least one metastable state, can produce amplification, hence lasing.

Acknowledgements

The author Madhura Murthy M N expresses her gratitude to the Department of Science and Technology, Govt. of India for motivating her to take up research and for the support and INSPIRE scholarship (SHE) awarded.

She acknowledges the guidance provided to her by Prof. Sarmistha Sahu, in every step of the experiment and preparing the presentation of the project.

References:

[1] Todd W. Neller, Clifton G.M. Presser, Ingrid Russell, Zdravko Markov. (2006)

[2] Arthur Murray & Ian Hart, Phys.Educ. **47,** 197. (2012).

[3] Ghatak Ajoy, *Optics* 3rd Edition, The McGraw-Hill Companies, pp 23.14-23.17.

[4] Sahu Sarmistha, Phys.Educ. **46**, 255-256 (2011).

[5] Sahu Sarmistha, Lab Experiments LE **39** ,11, 3 (2011)

The Architecture of the Standard Model

D. Banerjee and S. Sahoo[1]

Department of Physics, National Institute of Technology,

Durgapur–713209, West Bengal, India.

[1]E-mail: sukadevsahoo@yahoo.com

(*Submitted* dd-mm-yyyy)

Abstract

The standard model (SM) of elementary particles represents our present understanding of known fundamental fermions (quarks and leptons) and the forces between them. The electroweak theory and quantum chromodynamics have been unified into the SM, which is represented by the gauge group SU(3) X SU(2) X U(1). In this article, we discuss the architecture of the SM briefly.

1. Introduction

At the end of 20th century, the search for understanding the mysteries of matter and the forces which holds it together has created the theory of fundamental forces based on nonabelian gauge fields, physicists have named it the standard model (SM) of particle physics [1]. This model unifies the three fundamental forces: strong, electromagnetic and weak force. These forces (electromagnetic, weak and strong force) are mediated by the gauge bosons: the photon (γ); the W^+, W^- and Z^0 boson; and the gluons respectively. There are 6 types of quarks and 6 types of leptons in the SM. Quarks are called up (u), down (d), charm (c), strange (s), top (t) and bottom (b). Leptons are called electron (e^-), electron-neutrino (ν_e), muon (μ^-), muon-neutrino(ν_μ), tau (τ^-) and tau-neutrino(ν_τ). Fig.1 shows elementary particles in the SM and their mass, charge and spin. Higgs boson is expected to be the final particle of the SM. Recently, the discovery of this particle is confirmed in the large hadron collider (LHC) experiment at CERN, Geneva but till now it is not confirmed whether it is the SM Higgs boson or beyond the SM particle. This particle is responsible for the Higgs mechanism by which all particles acquire mass.

In the mid 19th century, Maxwell unified three different phenomena, i.e. electricity, magnetism and optics in one theory known as electromagnetic theory and established a set of four equations. These equations are gauge symmetric [2]. The classical form of Maxwell's equations (MEs) describes the EM field with its continuous energy distribution in space-time. But latter this theory was reinterpretended by Feynman, Schwinger and Tomonaga [3]. They replace the continuous energy distribution by discrete field quanta or discrete packets of energy. This is called field quantization and the revised electromagnetic theory was named as quantum electrodynamics (QED) which was the first quantum field theory (QFT) [2]. The relativistic approach of QFT can be the basis to describe all the fundamental interactions (except gravity). The fields involved in QFT are gauge fields. A gauge field is a physical field whose existence is inferred from "a principle of local gauge invariance", this means that the laws of physics are independent of local changes of the phase at all positions and in the universe. This is how the first QFT was coined. The quantum theory of strong force (colour force) is known as quantum chromodynamics (QCD). The two theories, the electroweak theory and the QCD, form the *standard model* (SM) of elementary particles.

Electricity + Magnetism + Optics = Electromagnetism

Electromagnetism + Quantum theory = QED

QED + Weak force = Electroweak theory

Strong force + Quantum theory = QCD

QCD + Electroweak theory = SM

In the next sections we explain the different QFTs starting from QED to QCD and the architecture of the SM briefly.

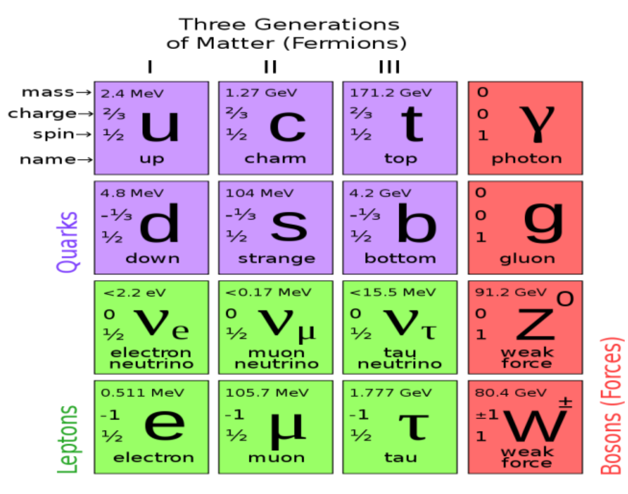


Fig. 1: **Elementary particles in standard model**

2. Quantum Electrodynamics

In particle physics, quantum electrodynamics (QED) is the relativistic quantum field theory of electrodynamics. In QED, each particle is described by a wavefunction which contains the information that can be used to calculate the probability of finding the particle at different locations. In QED, two charged particles interact by the exchange of electromagnetic vector boson of spin 0 known as photon. For example, let us consider the interaction between a proton and an electron. The proton emits a photon and the electron absorbs it. This is represented in the Feynman diagram shown in Fig. 2 [2]. Feynman diagram describes the particle interaction quantitatively and gives a diagrammatic visualization. Each external line represents a real particle and the Feynman rules define propagation of particle mathematically. The internal lines describe the intermediate particle which cannot be observed directly, but can only be observed in indirect decay process. QED is based on the symmetry group U(1) and the generator is electric charge Q. The strength of the interaction is determined by coupling constant. The coupling constant for the emission of an extra photon is 1/137 (known as fine structure constant). The Nobel Laureate Richard Feynman, one of the founding father of QED, has called QED as "*the jewel of physics*".

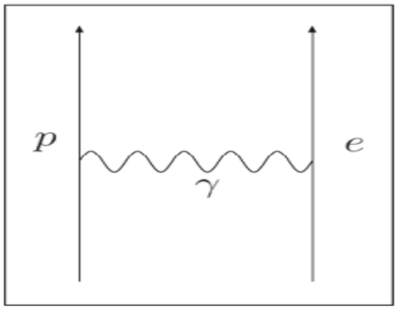


Fig. 2 [2]

3. Electroweak Theory

The weak force is a universal force which affects all particles. The intermediate vector bosons for weak interaction are W^+, W^- and Z^0 boson. Here W^+ and W^- have electric charge which carries out the charged current interactions and Z^0 boson is neutral which carries out the neutral current interactions. The W and Z^0 bosons carry weak isospin (T_3) and weak hypercharge (Y) and so they can interact with each other as well as with other particles. These self interactions are important in the SM. Fig. 3 [2] shows a weak interaction namely the decay of neutron into proton, electron and antineutrino (beta decay). Neutron and proton are the composites of three quarks udd and uud respectively. This figure shows that the d quark turns into a u quark by emitting the weak quantum W^- which turns into a pair of leptons (electron and antineutrino).

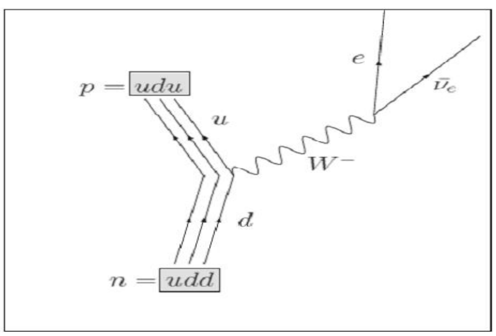


Fig. 3 [2]

The quantum electrodynamics and weak forces have been unified into EW dynamics as they are two facets of one entity called the electroweak (EW) theory. The EW interactions are mediated by four quanta called electroweak gauge boson (W^+, W^-, Z^0 and γ) and these gauge bosons were predicted by the EW theory. Fig. 4 [2] shows the Feynman diagram of the electromagnetic and weak interactions among the quarks and leptons mediate by electroweak quanta. From Fig. 4 we can see the exchange of W can generate a force between particles or can cause decay of a particle. But photon and Z^0 boson can only generate force among particles.

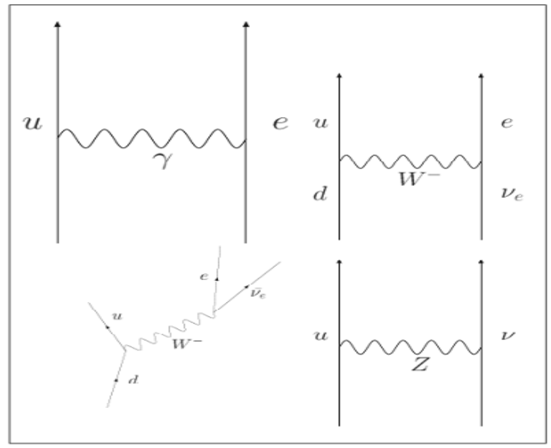


Fig. 4 [2]

4. Quantum Chromodynamics

A theory of quark interactions was constructed by analogy with QED. This theory postulated the existence of massless particles called gluons by which the quarks are held together and describes the strong interaction by using the colour charge of quarks and gluons. This theory is called as quantum chromodynamics (QCD) [4,5]. The theory of strong interaction involves the exchange of meson and between baryons where both mesons and baryons made of quarks [6]. The idea of colour was first put forward by Oscar Greenberg in 1965. Just to distinguish the colours, one normally calls the three colours red (R) , green (G) and blue (B) [5]. Let us consider the interaction between two quarks. In this case, the colour charge on each quark is changed after the interaction, whereas in QED the electric charge on the electron was not changed by the photon. The colour charge must be conserved at each vertex of the strong interaction. Hence this condition implies that each gluon must carry a colour/anticolour combination.

The gluon and the photon are very much different from each other in many ways [3].

Firstly, the photon has no electric charge but the gluons carries colour/anticolour charge. Then another point is that a photon cannot interact directly with another photon by the electromagnetic force because it is neutral. But a gluon-gluon interaction can occur via the strong force. This interaction is very important in LHC because the most interesting proton-proton collisions arise from gluon-gluon interactions. One more difference is that the coupling constant for emission of an extra gluon is much higher for the strong force than for photon emission in QED. For photon emission the coupling constant is 1/137 whereas the value of coupling constant for strong interaction is 1/10. An interesting nature of quarks is that, when they are very close together the force between them is very small and behave as free particles. This is known as *asymptotic freedom* and due to this property the quarks which are close together inside a proton barely interact with each other. But when a quark tries to move away from another quark in a bound system like proton the required energy to do this becomes infinitely large. This is the reason that no single quark has ever been observed. This property is called *quark confinement* [1-3, 5].

5. Architecture of the Standard Model

The standard model of particle physics is now known to be the most successful theory in describing almost all of known physics except gravity [2]. Six types of quarks: up (u), down (d), strong (s), charm (c), bottom (b) and top (t); six number of leptons: electron (e^-), electron-neutrino (ν_e), muon (μ^-), muon-neutrino(ν_μ), tau (τ^-), tau-neutrino(ν_τ); and Higgs boson with their antiparticles are the building blocks of the SM [1,8]. The SM is also known as *quark-lepton model*. In the SM the quarks and leptons are categorized in three generations (Fig. 1). The first generation particles are found in ordinary matters and they are the lightest among three generation particles, whereas the 2^{nd} and 3^{rd} generation particles are found in cosmic rays and can be reproduced in particle colliders [6]. The quarks have fractional charges. Each type of quark is a colour triplet and also carries electroweak charges, in particular electric charges + 2/3 for up-type quarks and -1/3 for down-type quarks. The electroweak theory and QCD are the two pillars of the SM. In the SM the fundamental forces are based on nonabelian gauge fields and the symmetry group SU(3)XSU(2)XU(1). This group has 8+3+1=12 generators with non-trivial commutator algebra. Not all the generators belong to SM commute with each other, so this theory is a nonabelian gauge theory [7]. SU(2)XU(1) symmetry group describes the electroweak interaction. One of the generator of SU(2) group is T_3 (the third component of isospin) and Y (hypercharge) is the generator of U(1). The electric charge Q, the generator of QED gauge group is given as $Q= T_3+ Y/2$. SU(3) is the colour group of the theory of strong interaction. Every symmetry group has one or more number of generators and each generator is associated with a vector boson or simply gauge boson with same quantum number. If the gauge symmetry is unbroken then vector boson will have vanishing mass. The gauge bosons are mediating particles for fundamental interactions. For example, photon is the vector boson which mediated the interaction between two charged particles in QED. Similarly, there are 8 gluons associated to the SU(3) colour generators, while for SU(2)X U(1) there are 4 gauge bosons : W^+, W^-, Z^0 and γ [9]. Hence, the SM contains total 12 mediating particles. Out of these 12 generators only the gluons and the photon are massless because the symmetry induced by the other 3 generators is actually spontaneously broken. The W^+, W^- and Z^0 are quiet heavier particles (M_W ~80.4 GeV, M_Z ~91.2 GeV). The mass problem is therefore primarily concerned with the force carriers of weak force. So, the completion of the SM required a way of including massive force carriers without breaking the symmetries that were crucial to its predictive power. It was found that via spontaneous symmetry breaking (SSB) [2] this could be achieved. There are many different examples of spontaneous symmetry breaking (SSB) in physics and in everyday life. If a pin is held vertically and downward pressure is applied the pin will be bent but the direction of this bend is random and not predictable. Starting from the top of a hill you have an almost infinite number of ways to walk down to ground level, where you have lower potential energy. The situation is symmetrical- we can choose any of the equivalent possible directions. When you choose one way down you have broken the symmetry. The idea of spantaneous breakdown of symmetry (SBS) in particle physics originates from Nambu although he applied it in a different context [3]. In a gauge theory like the SM the SSB is realised by the Higgs mechanism. Higgs mechanism postulates the existence of a universal field called the Higgs field and this field gives masses to W and Z^0 and also gives masses to all the femions. The variation of Higgs field with potential energy is shown in Fig. 5 [3]. The vertical axis is the magnitude of potential energy and the horizontal axes represent the magnitude of the Higgs field. When the Higgs field is zero (point A) the potential energy is maximum. When the spontaneous symmetry breaking occurs, the potential energy is reduced to one of the minimum points around the valley (point B) and the Higgs field becomes nonzero. The radius of the bottom of the valley is related to the magnitude of the Higgs field.

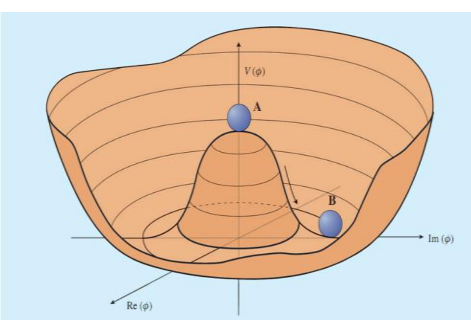


Fig. 5 [3]

According to the scientists, when the universe was created at that moment there was no Higgs field, so all the particles were massless and travelled at the speed of light. Gradually as the universe cooled the potential energy of the universe decreased and the Higgs field became non- zero. In this process the Higgs field breaks the symmetry. Then the particles which couples with the Higgs field acuires mass. But there is an important byproduct of the Higgs mechanism: a massive spin zero boson, called the Higgs boson, must exist as a relic of the original Higgs field. Although it is theoretically predicted almost half a century ago, scientists are still trying to confirm the existence of it and to find its properties. The mass of Higgs boson is not specified in the SM. From the LHC experiment, the mass of this particle is found to be 125.3± 0.4 (stat) ± 0.5 (syst) GeV (CMS) [10] and 125.5 ± 0.2 (stat.)$^{+0.5}_{-0.6}$ (syst.) (ATLAS) [11]. At the Tevatron, the CDF and D0 Collaborations have observed the mass of Higgs boson in the range 115 to 140 GeV [12]. Their observed signals are consistent with a standard model Higgs boson with a mass of 125 GeV. Recently, on 14th March, 2013, the 'Moriond' conference held in Italy, has confirmed the existence of Higgs particle [13]. But till now it is not confirmed whether it is the SM Higgs boson or beyond the SM particle.

6. Further Discussions

The SM is the most successful theory for describing the phenomena of fundamental interactions of elementary particles. In spite of this fact this model has several drawbacks; some of them are discussed below.

Gravity, one of the fundamental forces, is not included in the SM. So, the physicists are not satisfied with the unification at this stage. The ultimate goal of particle physics is to construct a unified theory that would reveal how all observed particles and forces are just different manifestations of a single underlying system, which can be expressed within a common mathematical framework [14]. The gravitational force is based on the Einstein's general theory of relativity. This theory is based on the principles of classical mechanics and not of quantum mechanics. Since other forces in nature obeys the rules of quantum mechanics, any theory that attempts to explain gravity as well as the other forces of nature must satisfy both gravity and quantum mechanics. General theory of relativity describes gravity at large distances whereas quantum mechanics works at very small distances. In quantum theory the range of a force is inversely proportional to the mass of the quantum that is exchanged. Since gravity has infinite range, quantum theory of gravity (if it is constructed) will have its quantum, called graviton, with zero mass and spin-2. But till now Einstein's general relativity has resisted all attempts at being combined with the quantum world as we do not have a quantum theory for spin-2 particle. Hence, quantum gravity had become the most fundamental problem of physics at the end of the twentieth century. This is the reason for the rise of string theory, for it promises to be a theory of quantum gravity. It is seen that string theory [15–18] works very well at large distances where gravity becomes important as well as at small distances where quantum mechanics is important. In the standard model the elementary particles are mathematical points. But in string theory, instead of many types of elementary point-like-particles, we assume that in nature there is a single variety of one-dimensional fundamental objects known as strings [15]. It is not made up of anything but other things are made up of it. Like musical strings, this basic string can vibrate, and each vibrational mode can be viewed as a point-like elementary particle, just as the modes of a musical string are perceived as distinct notes. String theory [19,20] provides a framework to address some fundamental issues in cosmology and elementary particle physics.

The number of parameters in the SM is too large to regard it as a fundamental theory and also there is no such underlying principle to choose the parameters. The Lagrangian of the SM contains many arbitrary parameters such as the three gauge couplings corresponds to three parts of the gauge group, flavour mixing parameters (CKM mixing angles and phases), coupling constants of Higgs self interaction, and Higgs mass parameter. The three gauge groups involved in the SM were seemed to be unrelated, which means that the formal unification of strong and electroweak interaction is also a question.

Neutrinos are massless in the SM, because neutrinos do not couple with Higgs field. But, about 15 years ago, experimenters discovered that neutrinos do have tiny masses and this has been hailed as a great discovery since this may show us how to go beyond the SM. The dominance of baryonic matter over antimatter is not explained in the SM. The problem of CP-violation is not well understood including CP-violation in strong interactions. The SM does not answer why there are only three generations not more than that.

According to the standard model of cosmology, the universe contains 4 % visible, 23 % dark matter and 73 % dark energy. These measurements rely on the validity of the hot big bang model, general relativity and the cosmological principle (that the universe is uniform on the largest scales). But the SM says nothing about where these dark matter and dark energy came from. What is the nature of dark matter and dark energy?

Apart from the above discrepancies of the SM, there are some more problems which have to be taken care. Such as, the SM can't explain the Higgs interaction's of different specific couplings to different particles, it is not valid above Planck Scale ($\sim 10^{19}$ GeV). Both the srting and the extra curled-up dimensions will be revealed only when we can access Planck's length scale ($\sim 10^{-33}$cm). But so far we have reached only 10^{-17}cm. So, we have a long way to go. Quantum gravity is an important part to explain the today's physics. Apart from string theory there are other approaches to quantum gravity. We hope the future discoveries from LHC will show us the right choice of a theory which would be able to explain all the four fundamental forces exist in nature.

References

1. D. Griffiths, *Introduction to Elementary Particles*, John Wiley and Sons, Singapore (1987).
2. G. Rajasekaran, *Resonance*, **17**, 956 (2012).
3. K. E. Johansson and P.M. Watkins, *Physics Education*, **48**, 105 (2013).
4. S. Sahoo and R. K. Agarwalla, *IAPT Bulletin*, **25**, 293 (2008)
5. R. M. Godbole, *Fundamental particles and their interactions*, in Horizons of Physics Vol.I Ed. A. W. Joshi, Wiley Eastern Ltd., Delhi (1989).
6. S. Sahoo, *Physics Education*, **22**, 85 (2005) (IAPT publishing).
7. A. Sen, Resonance, **5**, 4 (2000).
8. S. Sahoo, *IAPT Bulletin*, **25**, 228 (2008).
9. Guido Altarelli, arXiv:1303.2842 [hep-ph] (2013).
10. S. Chatrchyan *et al.* (CMS collaboration), *Phys. Lett. B* **716**, 30 (2012) [arXiv: 1207.7235 [hep-ex]].
11. S. M. Consonni (on behalf of the ATLAS Collaboration), arXiv:1305.3315 [hep-ex] (2013).
12. T. Aaltonen et al., arXiv:1303.6346 [hep-ex] (2013).
13. H. Dreiner, *Nature Physics*, **9**, 268 (2013).
14. M. M. Waldrop, *Nature*, **471**, 286 (2011).
15. J. H. Schwarz, *Curr. Sci.*, **81**, 1547 (2001).
16. S. R. Wadia, *Curr. Sci.*, **95**, 1252 (2008).
17. A. Sen, *Resonance*, **10**, 86 (2005).
18. S. Sahoo, *Eur. J. Phys.* **30,** 901 (2009) [arXiv:1209.5498 [physics.gen-ph].
19. M. Dine, *Physics Today*, **60**(12), 33 (2007).
20. M. Dine, *Supersymmetry and String Theory: Beyond the Standard Model*, Cambridge University Press, New York (2007).

A Group Theoretical Analysis Of Constituent Gluons In Scalar Glueballs

Tapashi Das[1] and D.K. Choudhury[1,2,3]

[1]Department of Physics
Gauhati University
Guwahati 781014, India
t4tapashi@gmail.com

[2]Physics Academy of the North East
Guwahati 781014, India

[3]Centre for Theoretical Studies
Pandu College
Guwahati 781012, India

(Submitted 20-03-2013)

Abstract

Young Tableau in Group theory is a method of determining number of irreducible representations of direct products in SU(N) groups, while glueballs are predicted hadronic states in QCD consisting only of Gluons without quarks and antiquarks. We present results of our work using it in determining the number of constituent gluons in low lying scalar glueballs.

1 Introduction

Quantum Chromodynamics (QCD) is the theory of the hadronic interactions. In QCD, the gluons are massless and have spin-1. Gluons themselves carry color charge, and so they can interact with other gluons [1]. In particle Physics, a glueball is a hypothetical composite particle. It consists solely of gluons, without valence quarks. Such a state is possible because gluons carry color charge and experience the strong interaction. Glueballs have not been an easy subject to study due to the lack of phenomenological support and therefore much debate has been associated with their properties. Glueballs are extremely difficult to identify in particle accelerators, because they mix with ordinary meson states. Glueball is mainly produced in radiative J/ψ decay and $p\bar{p}$ annihilation. Theoretical calculations based on nonperturbative methods like Lattice QCD and QCD sum rules agree that the lightest glueball should be a scalar resonance ($J^{pc} = 0^{++}$) with a mass range 1400-1800 MeV^6 followed by a tensor (2^{++}) and a pseudoscalar (0^{-+}) glueball in the 2000 -2500 MeV mass region [2]. So far,there has been 20 years of intensive experimental search towards glueballs. Unfortunately, no definite answer to the question whether a glueball has been observed or not can be given. It is to be mentioned that although gluons are color octet, glueballs are color singlet. So a single gluon cannot be a glueball, but a gluelump [3]. The forthcoming experiment FAIR has PANDA as detector [4], specifically designed to detect glueballs, hybrid mesons and charmonium spectroscopy. Theoretical calculations show that glueballs should exist at energy ranges accessible with current collider technology. However, due to the aforementioned difficulty, they have so far not been observed and identified with certainty. But still it is a worthwhile theoretical pursuit.

In the present paper, we report the results for scalar glueballs using Young Tableau. In the previous publications [2,5], we have presented the corresponding results for pentaquark multiplets.

In section 2, we outline the formalism; section 3 the results; section 4 the experimental status of glueballs; whether section 5 is devoted to conclusion.

2 Formalism

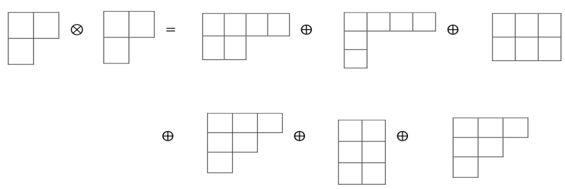
In this section we outline the method how group theoretical tool of Young Tableau of $SU(3)_c$ can be used to find the maximum number of constituent gluons in an experimentally observed scalar glueball.

In Mathematics, a Young tableau is a combinational object useful in representation theory. Young tableaux were introduced by Alfred Young, a mathematician at Cambridge University, in 1900. By taking the direct product of irreducible representations we can generate the representations of higher dimensions. These representations are however reducible. Young tableau gives a definite way of reducing it to direct sum of various irreducible representations [6,7]. The theory in detail is discussed in ref. [6,7].

However this tool can only be used for low lying scalar C= +1 glueballs. Because only for scalar Glueball l=0 and s=0; J=0. Similarly in case of scalar Glueball there will not be any additional angular momentum multiplicity.

3 Results

Here we show the product result for $(8_c \times 8_c)$ in $SU(3)_c$:



$$= 27 + 10 + \bar{10} + 8 + 1 + 8$$

Figure 1: Results for glueball consist of two Gluons

This shows two color-octet gluons can form a color singlet. Thus if only one glueball is observed, the glueball must be composed two gluons. The lightest scalar glueball mass is in the range of 1500-1750 MeV [8]. Similarly, from the bound state of three or more gluons the corresponding representations can be found out (Table-1).

Table-1 shows the number of color singlet glueballs and the number of constituent gluons:

Table 1: Number of color singlets corresponding to the maximum no. of constituent gluons

Direct product	Direct sum	Number of Color Singlet
$8_c \times 8_c$	$\underline{1}+\underline{8}+\underline{8}+\underline{10}+\underline{\bar{1}0}+\underline{27}$	1
$8_c \times 8_c \times 8_c$	$2(\underline{1})+8(\underline{8})+4(\underline{10})$ $+4(\underline{\bar{1}0})+6(\underline{\bar{2}7})+4(\underline{35})+\underline{64}$	2
$8_c \times 8_c \times 8_c \times 8_c$	$8(\underline{1})+32(\underline{8})+22(\underline{10})+$ $18(\underline{\bar{1}0})+33(\underline{27})+4(\underline{28})+30(\underline{35})$ $+12(\underline{64})+6(\underline{81})+\underline{125}$	8
$8_c \times 8_c \times 8_c \times 8_c \times 8_c$	$32(\underline{1})+145(\underline{8})+117(\underline{10})+$ $83(\underline{\bar{1}0})+180(\underline{27})+40(\underline{28})$ $+200(\underline{35})+94(\underline{64})+10(\underline{80})$ $+72(\underline{81})+20(\underline{125})+8(\underline{154})+$ $\underline{216}$	32
$8_c \times 8_c \times 8_c \times 8_c \times 8_c \times 8_c$	$145(\underline{1})+702(\underline{8})+642(\underline{10})+$ $408(\underline{\bar{1}0})+999(\underline{27})+322(\underline{28})$ $+1260(\underline{35})+10(\underline{55})+660(\underline{64})$ $+140(\underline{80})+630(\underline{81})+215(\underline{125})$ $+140(\underline{154})+18(\underline{162})+30(\underline{216})$ $+10(\underline{260})+\underline{343}$	145
$8_c \times 8_c \times 8_c \times 8_c \times 8_c \times 8_c \times 8_c$	$702(\underline{1})+3598(\underline{8})+3603(\underline{10})+$ $2109(\underline{\bar{1}0})+5670(\underline{27})+7840(\underline{35})$ $+2352(\underline{28})+168(\underline{55})+4424(\underline{64})$ $+1400(\underline{80})+4872(\underline{81})+1890(\underline{125})$ $+1568(\underline{154})+336(\underline{162})+426(\underline{216})$ $+239(\underline{260})+41(\underline{343})+28(\underline{280})$ $+12(\underline{405})+\underline{512}+\underline{273}+\underline{330}$	702

It shows that the Young Tableau calculation gives us the possibility to infer the maximum number of constituent gluons from the given number of experimentally observed scalar glueballs, as shown in table-2 upto multiplicity 702.

Table-2 shows maximum number of constituent gluons in such Glueball muliplicity.

Table 2: Possible number of constituent gluons corresponding to the number of scalar glueballs observed

No. of scalar glueballs observed	Possible No. of constituent gluons
1	2
2-7	2,3
8-31	2,3,4
32-144	2,3,4,5
145-701	2,3,4,5,6
702	2,3,4,5,6,7

4 Experimental status of Glueballs

Good evidence exists for a scalar glueball which is mixed with nereby mesons, but a full understanding is still missing. Evidence for tensor and pseudoscalar glueballs are weak at best [9]. Glueballs should, e.g., be produced preferentially in so-called gluon-rich processes [9,10]:

(i) **Radiative J/ψ decay**: In the most decays, the J/ψ undergoes a transition into 3 gluons which then convert into hadrons. But the J/ψ can also decay into 2 gluons and a photon

$$\mathrm{J}/\psi \longrightarrow \gamma gg \longrightarrow \gamma G \qquad [9]$$

The photons can be detected, the two gluons interact and must form glueballs - if they exists [10].

(ii) $p\bar{p}$ **annihilation**: In $p\bar{p}$ annihilation, quark-antiquark pairs annihilate into gluons, they interact and may form glueballs [10].

(iii) **Central production**: In central production two hadrons pass by each other 'nearly untouched' and are scattered diffractively in forward direction. The valence quarks are exchanged. The process is often called Pomeron-Pomeron scattering. The absence of valence quarks in the production process makes central production a good place to search for glueballs [10].

There are many new experiments planned, e.g. the PANDA Experiment at GSI in Germany [11], BES III at BEPC II in Bejing [12], the GlueX Experiment at Jefferson Laboratory in the USA [13], ALICE at CERN [14][15] which will provide us more data on this. On the one hand, some experimental glueball candidates are currently known. Most of them are scalar, such as the $a_0(980)$, $f_0(980)$, $f_0(1500)$, $f_0(1710), ...$ but no definitive conclusions can be drawn concerning the nature of these states [16] nor its exact multiplicity.

5 Conclusion

We have shown how one can use Young Tableau to infer the number of costituent gluons from the multiplicity of low lying glueballs. In lattice QCD calculation, low lying C= +1 glueballs are identified with two gluon states (or at least with hadrons in which two gluon components widely dominent), while no light C=-1 (l+s= odd) and heavy scalar $J^{pc} = 0^{+-}$ glueballs are seen as light four gluon states. The group theoretical approach cannot make distinction among the glueballs of different number of constituent gluons. Only constraint is that the glueballs must have total orbital momentum and total spins of the n constituent gluons be seperately 0, so that angular momentum does not have role in its spectroscopy. The method however falls short of accomodating gluelump [3] having single constituent gluon with additional gluon field.

Acknowledgement

Most of the work was completed when one of the authors (T.Das) was completing her M.Sc. dissertation in the department of Physics, Gauhati University. She gratefully acknowledges support given to her by the department. The authors also thank Dr Dipali Das Kar, Head of the Department of Physics, Pandu College, Guwahati for her kind support.

References

[1] Quarks and Leptons: An Introductory Course in Modern Particle Physics: Francis Halzen and Alan D.Martin. ISBN 0-471-8874I-2 (1984).Page-206

[2] Sumita Kumari Sarma and D. K. Choudhury, Physics Education 24(2009)267 (ISSN 0970-5953)

[3] 'Constituent gluon interpretation of glueballs and gluelumps': Nicolas Boulanger,Fabien Buisseret, Vincent Mathieu and Claude Semay arXiv:0806.3174v2 [hep-ph] 18 Aug 2008

[4] Hadron Physics with Anti-protons: The PANDA Experiment at FAIR, arXiv:0711.1598v1 [nucl-ex] 10 Nov 2007

[5] Pentaquarks and Glueballs by Sumita Kumari Sarma and D.K. Choudhury (LAP Lambart Academic Publication, Germany, 2010), ISBN 978-3-8383-9984-3

[6] Bruce E. Sagan. The Symmetric Group. Springer, 2001, ISBN 0387950672

[7] Yong, Alexander (February 2007). "What is...a Young Tableau?" (PDF). Notices of the American Mathematical Society 54 (2): pp.240241.

[8] Glueball States in a Constituent Gluon Model : Wei-Shu Hou, Ching-Shan Luo and Gwo-Guang Wong
arXiv: hep-ph/0101146v3, 1 Mar 2001

[9] The Experimental Status of Glueballs: V. Crede and C. A. Meyer

[10] Glueballs, Hybrids, Pentaquarks: Introduction to Hadron Spectroscopy and Review of Selected Topics :Eberhard Klempt
arXiv: hep-ph/ 0404270v1 29 Apr 2004

[11] D. Bettoni, J. Phys. Conf. Ser. 9 (2005) 309.

[12] M. S. Chanowitz, Int. J. Mod. Phys. A 21 (2006) 5535.

[13] D. S. Carman, AIP Conf. Proc. 814 (2006) 173.

[14] B. Alessandro et al. [ALICE Collaboration], J. Phys. G 32 (2006) 1295.

[15] A candidate for the scalar glueball operator within the Gribov-Zwanziger framework: Nele Vandersickel, David Dudal, Henri Verschelde
arXiv:0910.2653v1 [hep-th]14 Oct 2009

[16] Gluons in glueballs: Spin or helicity? Vincent Mathieu, Fabien Buisseret,and Claude Semay.

The RC circuit experiment

S. Ganci[1]

[1]Studio di Catalogazione e Conservazione Strumenti Scientifici,

Casarza Ligure –GE- 16030, Italy.

(*Submitted* 13-06-2013)

Abstract

The RC circuit experiment reveals critical aspects about the systematic perturbation introduced in a circuit by an instrument. A pedagogical value experiment is suggested.

1. Introduction

The experiment on the charging and discharging of a capacitor through a resistor has been, to date, been presented by the author using a square wave generator and an oscilloscope. The Analog oscilloscopes employed and the low time constant of the circuits did not pose particular critical aspects. A typical setup was an RC circuit whereby R = 100 Ω and C = 1 μF, a square wave generator with a frequency of 500 Hz and time base 0.2 ms/cm on the oscilloscope. It was only towards the end of his teaching career, that the author discovered and appreciated the versatility of dataloggers. In re-proposing the RC experiment the role of the meter resistance becomes relevant, as highlighted in Literature [1-4].

This paper describes a critical presentation of the experiment in order to concretely introduce the concept that every instrument inserted in a circuit can introduce systematic perturbations on the measurement.

2. The experiment.

Two RC circuits with the same *time constant* τ =RC are mounted on two Plexiglas sheets visually illustrating the schematics and components to the students. One of the two circuits is made using a 1 MΩ resistor and a 1 μF polyester capacitor and the other, using a 100 Ω resistor and a 10000 μF electrolytic capacitor. A switch connects the RC series to the battery (charging phase) or shorting C on R (discharging phase). A packet of four AA rechargeable batteries in series simplifies the setting up of the experiment. In both cases the time constant is one second, which is quite a long time when compared with the experiment on the oscilloscope. The experiment can be presented starting with the first circuit and a datalogger. Using a DS1M12 datalogger, the charge and the discharge of the capacitor are easily shown. The Datalogger DS1M12 has two input channels and its software allows the use of one or both channels, as an oscilloscope or as datalogger. In this experiment, a single channel is used. The output of the datalogger enters the USB port of a Notebook containing Software and the drivers to the OS system used. In our case, Windows 7 and a Notebook with a well visible screen was used for a demonstration experiment. The Instructor can show the voltage across the capacitor or the resistor. A screen snapshot can then be printed, as a record for the students, another advantage of dataloggers over the use of an ordinary CRO for this experiment. Indeed, the use of a datalogger for a classroom experiment means that the instructor need only choose the sample interval and the interval on the *x* axis in order to have the charge and discharge process in a single window. The Instructor may call student's attention on the fact that any instrument in a circuit introduces a systematic perturbation (the real circuit is shown in Fig. 1). With our datalogger a sample interval of 200 ms and a *x* width of 20 s (or more) gives satisfactory results as shown in Figs. 2 and 3 and the above parameters appear in the figures. Another feature of the software is that the *x* and *y* parameters can be changed even after file has been saved. An attentive student can recognize that the voltage drop across the capacitor in the first circuit (R = 100 Ω, C = 10000 μF) is lower than the value V_o measured across the battery (V_o = 4.85 V) as shown in Fig. 4. Ask your students why this happens. The experiment also shows how a capacitor does not have a perfect electrical isolation. Using the second circuit (R = 1 MΩ and C = 1μF) the data logger shows a voltage drop at the ends of C which is approximately *half* of the voltage across the battery. There are two reasons. The ohmic resistance across the 1μF capacitor *is not negligible* and measured with an Ohm-meter gives a value around 1 MΩ while in our 10000 μF capacitor, the resistance is greater than 20 MΩ. It should not be difficult for the student to explain why the voltage across the capacitor is about half that across the battery. This analysis must take into account the internal resistance of the voltmeter used i.e. the input impedance of our datalogger. Following these qualitative observations the true circuit equation can be written. Using Kirchhoff's laws applied to the equivalent circuit in Fig. 1, the resistance of our voltmeter R_v (datalogging input) is taken into account [4].

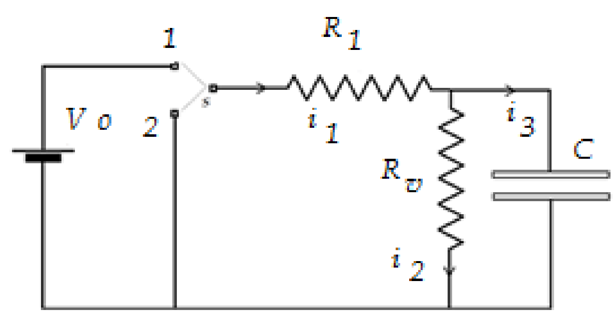


Fig. 1. The experimental setup. A little notebook and a low cost DS1M12 oscilloscope/datalogger is used. As D. C. source, four AA 1.2 V rechargeable batteries in series were used. The only practical caution is the use of a good quality switch.

Still applying Kirchhoff laws and referring to the notations in Fig. 2 we have

$$\begin{cases} i_1 = i_2 + i_3 \\ V_o - R_1 i_1 - R_v i_2 = 0 \\ R_v i_2 - \dfrac{1}{C}\displaystyle\int_0^t i_3 dt = 0 \end{cases} \tag{1}$$

Substituting the first equation in the second and in the third equation we find the relation

$$\frac{R_v}{R_1+R_v}V_o-\frac{R_1R_v}{R_1+R_v}i_3=\frac{1}{C}\int_0^t i_3 dt \qquad (2)$$

Differentiating this equation it follows a linear equation

$$\frac{R_1R_v}{R_1+R_v}\frac{di_3}{dt}=\frac{1}{C}i_3 \qquad (3)$$

Whose solution

$$i_3=A\exp\left(-\frac{t}{\tau}\right) \qquad (4)$$

where A is a constant to be determined from the condition at the time t = 0 and

$$\tau=\frac{R_1R_v}{R_1+R_v}C \qquad (5)$$

In the discharging process $A = V_o/R_p$ where R_p is

$$R_p=\frac{R_1R_v}{R_1+R_v}$$

In the charging process a particular integral must be added so that, in the charging process the equation is, as follows:

$$i_3=\frac{V_o}{R_p}\left[1-A\exp\left(-\frac{t}{\tau}\right)\right] \qquad (6)$$

Aside from the analytical treatment, an interesting feature of a datalogger is an experimental estimate of the characteristic time. This can be made printing the screen after an opportune choice of *x* and *y* parameters has been made. A young student can have fun with DS1M12 software and its features, an old teacher, such as the Author, can print the diagram on paper and evaluate the characteristic time with a pencil, having previously established an accurate scale factor on the print. The estimated measurements are: First circuit: data in Fig. 2. R =1 MΩ rand C = 1 μF

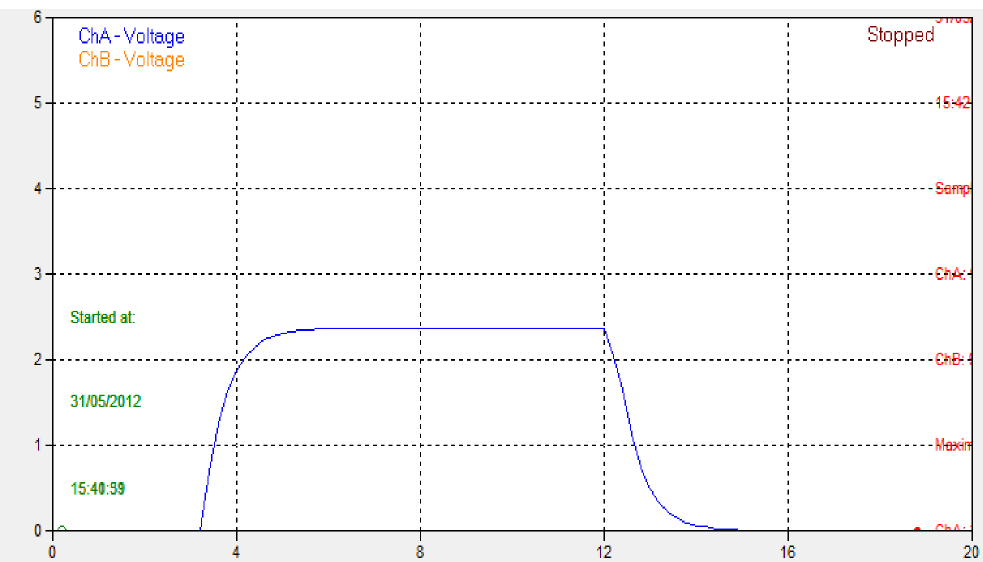


Fig. 2. Charging and discharging of the circuit having R =1 MΩ rand C = 1μF.

where the uncertainty on τ was estimated using "1 mm resolution" on the rule used in printed diagram. Supposing that R and C are without uncertainties, Equation 5 gives $R_v \approx 1.5$ MΩ in the correct magnitude order with a measurement of the capacitor *ohmic loss* through an Ohmeter (R = 1 MΩ ± 0.05 MΩ).

Second circuit: data in Fig. 3. R =100 Ω and C = 10000 μF: τ ≈ 1.1 s ± 0.1 s

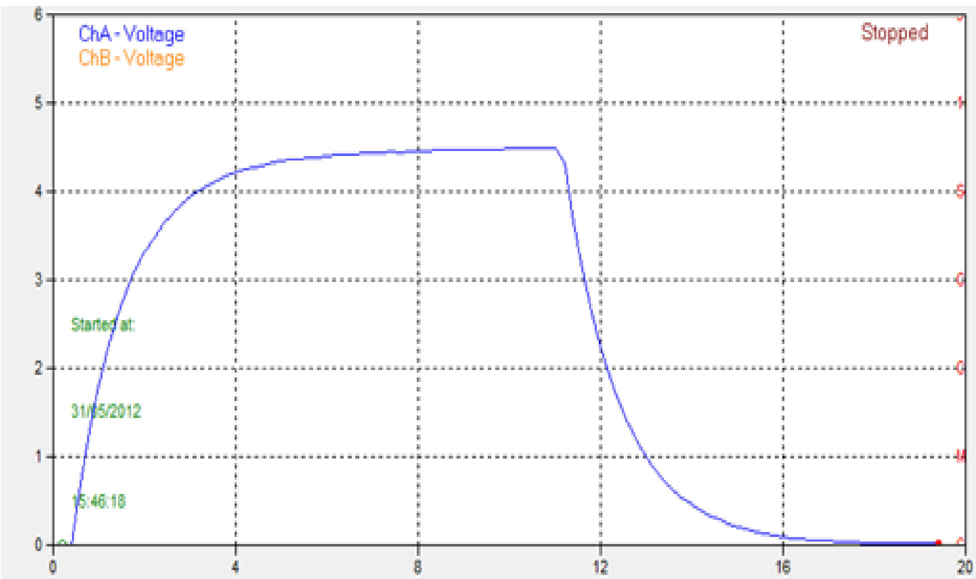


Fig. 3. Charging and discharging of the circuit having R =1 MΩ rand C = 1μF.

Within uncertainties evaluated, no perturbations appeared in this measurement. No ohmic loss minor of 20 MΩ is detected by a good ohmmeter. Student can use Eq. (5) in order to evaluate R_v finding a negative value if this slightly over 1 s value is inserted in Eq. (5). We leave to an attentive student the following matter: "why equilibrium voltage at the ends of the capacitor in Fig. 3 is minor than voltage across the battery measured by the datalogger and shown in Fig. 4?

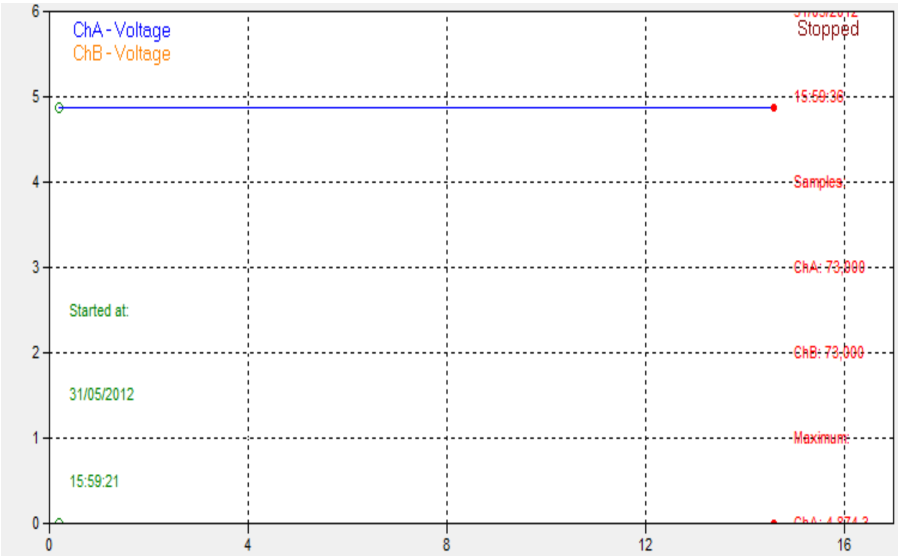


Fig. 4. Measured voltage across the battery. Datalogger and a good voltmeter gives values in agreement within the uncertainty of the voltmeter used (a good electronic instrument)

Can we *estimate* the impedance of our datalogger at work? Or in the diagram in Fig. 3 a long time were required for a *complete* charge?

3. Conclusion.

A matter usually emphasized as a good use of the mathematical language in General Physics shows surprising matter of discussion if a teacher shows it in real time with the classroom lesson. As shown, the experiment appears more interesting because quantitative estimate of time constant becomes possible. Nowadays students are more experienced with notebooks and Lab software than with classic Lab instruments, a good link between an old experienced teacher and a young attentive student.

References:

[1] H. T. Wood, Phys. Teach., **31**, 372-373, (1993).

[2] F. X. Hart, Phys. Teach., **38**, 176-177, (2000).

[3] Y. Kraftmakher, *Experiments and Demonstrations in Physics*, (World Scientific Company, Singapore, 2007), pp 16 -17.

[4] J. Priest, Phys. Teach., **41**, 40-41, 2003).

Concerning an Approach of the Implementation of Environmental Education in the Physics Teaching System

Kamo I. Atayan

International Education Centre
Yerevan State University
55[a] Tsarav Aghbyur st. Yerevan 0052, Armenia.

(*Submitted* 21-03-2013)

Abstract
This article presents a new approach of the implementation of environmental education in physics teaching system, which is particularly based on the use of integrative thematic and special physic- ecological modules. More than 30 similar integrative modules, which are included in the physics teaching course whenever required, have been formed complying with the school curriculum.

1. Introduction

At the beginning of the 21st century, the issue of the implementation of the environmental education has become more urgent and required. It is actually considered to be one of the stable bases for the further development of the civilization. Naturally, there exist various methods and approaches to the solution of this major problem in different educational systems, which particularly differ from each other in the level of their productivity. We can present the suggested approach in the following way: the sustained period of the environmental permanent education in school began with the course “The Basics of Ecology” which induces the learners to obtain environmental knowledge.

It is assumed that in secondary school, the teaching of the physics course begins with the textbooks and curriculum containing the so-called ecologized educational material. In high school, the physics is studied in a more advanced level by students with the natural sciences bias and more superficially by students having humanitarian bias. In this level of education, environmental education is integrated into the physics course, which does not decrease in any way the level of students' interest shown to the physics, on the contrary, according to the research results, it rises.

In this educational level, students of the natural sciences bias study physics by more advanced and ecologized textbooks, where integrative physics syllabus with ecological potential is highlighted and represented appropriately. The same process is carried out in the humanitarian bias of physics syllabus. For all secondary and high school groups studying physics, physico-ecological selective courses are intended, such as "Physics, Technology and Ecology", “Physics of Biosphere", "Technique and environment", etc. In this approach, the efficiency of environmental education increases, as at the end of the secondary and high school learning, when the study of the main natural sciences subjects is over, students study in different levels "Country is my home" selected natural sciences-ecological course. The mentioned approach has its own peculiarities, the performance algorithm and elements. It is clear that the mentioned issues are not possible to be discussed in appropriate depth in one article. That is why in this article it is essential, in the framework of the approach, to introduce block-modular integrative teaching problem of the implementation of environmental education in the physics teaching system. In the teaching process, the block modular system is being applied for a relatively long time. Nowadays, when the problem of the integration of the environmental education into the teaching process along with all the other subjects and courses at schools is aroused, it means to use all the opportunities of the interdisciplinary relations of the subject and ecology and the role of the mentioned method increases.

The generalization and systematization of the physical knowledge, teaching of the skill of its substantive acquirement, the performance of various methods and forms of teaching are the main characteristics which are typical of the block-modular approach of the informative and cognitive teaching, suggested and productive in the physics teaching system.

2. Thematic and special modules

It is well-known that the module teaching is based on the following principles: modulation, structuring, mobility, flexibility and equality. In the system of physics teaching, one of the main and active components of the mechanism of our newly suggested approach is the comprehension, introduction and application of the concept of the so called thematic and main integrative physico-ecological modules.

The thematic physico-ecological integrative module appears to be, from didactic and psychological point of view, a well-grounded and acceptable combination of the main educational material, concerning the particular theme of physics, and its adjacent subject, ecology. The modular material formed on the integrated and interdisciplinary relations, after being perceived by the learners, creates a complete notion of both of the issue discussed and the deep comprehension process of the separate component subjects.

The special integrative module is physical by nature and ecological by its expression, it is a complex of close and didactic-psychological conditions meeting problems and questions, based and newly integrated into the involved special ecological theme and the concrete section of the physics course study. It is provided in the form of the generalized course or lecture to the learners of the natural sciences / physic mathematical/ or the humanitarian directions in the secondary and high schools.

The experiments have shown that during the process of working with physico-ecological thematic and teaching informative modules the cognitive style of thinking of the learners must be necessarily considered. The latter appears to be an individual, unique processing way or form of the environmental teaching information (block) provided to the learners, the inseparable elements of which are the perception, analyzing and evaluation of structuring.

The sequence or algorithm of the physico-ecological special and thematic informative teaching module construction steps could be described as follows:

Step1. Formulation of the theoretical educational content that becomes the block-module.

Step2. The formation of the algorithmic description of the necessary skills and capabilities of the block

Generally, one of the peculiarities of the block-modular education is the determination of the steps undertaken for reaching the educational goal. The goal is clear; the educational material is to be perceived totally by the learners. For the realization of the block modular education, it is essential to have a so-called technological map. It is a definite form of planning the educational material and the scheme of its module structuring is as follows:

- The exact module naming
- The suggestion of the integrative didactic goal
- The presentation of the learners' educational process curriculum
- The existence of the academic informative base

The suggested physico-ecological thematic and special educational modules, by the teacher's discretion and methodological recommendation, are attached to the components of the abovementioned plan, without breaking the structure of that plan. Thus, thematic physico-ecological modules are expected to be applied at the end of the educational material. As it has already been mentioned there is an ecologization process of educational content material of physics; accordingly , all the sections of

school (and not only) physics course and almost all the corresponding themes, which have a definite ecological potential are ecologized by reasonably integrated ecological ideas, terms, conceptions in the physics content course.
During the educational process, while applying the suggested physico-ecological modular teaching method, actually both the past material and a physico-ecological integrative teaching course in the whole educational module are given, which have a complementary, developing and generalizing property in terms of content.
Apart from this, if the suggested method contains a special informative teaching module, it turns out that basically in the teaching process of the following physics theme, the learners have an opportunity to review the following physico-ecological theme three times.
We must also mention that it is not excluded that the same learners could be related to the similar physico-ecological subject, if they are included in the selected physico-ecological courses as well.
We must also note that the name of the informative teaching module corresponds to the title of the following physics school course.
The current and final checking of the learners' knowledge, capabilities and skills is highlighted particularly by the integrative modular method of preparing the learners for the physico-ecological education.
The current and final checkings are organized both orally and by thematic and generalization test checking. The presentation of the new course is realized in different forms; lecturing, retelling, slide show and so on. If we admit that nowadays the center of the teaching process is aimed to pass almost completely to the learner, it would be regarded in the educational process as a transition of the passive perception of the knowledge to the active process of the formation of necessary skills.
In this context, the suggested informative teaching modular technologies are sufficiently productive and provide high ECE. We suggest the following series for special and thematic modules.

Table 1

N/N	Thematic physico-ecological micro-module	Special physico-ecological micro-module	
TPEM-1	Concerning the connection of physics and ecology	Ecological concepts applied in the physics course	SPEM-1
TPEM-2	The ecological element of the section "Kinematics"	The physical sources of the pollution of the environment	SPEM -2
TPEM -3	The ecological element of the section "Dynamics "	The greenhouse effect. The issues of the global warming	SPEM -3
TPEM -4	The ecological element of the section "Law of Conservation in Mechanics"	The physical basics of the ecological observation of the environment	SPEM -4
TPEM -5	Noise and vibration pollution sources of the environment	The issue of the ozone from the point of view of physics	SPEM -5
TPEM -6	The ecological element of the theme ‘’Oscillations and Waves (mechanical)’’	Scientific and technological progress and ecological issues	SPEM -6
TPEM-7	The ecological potential of the theme "Transitional Phenomena"	Energetics and ecology	SPEM -7

TPEM -8	The ecological element of the section "Basics of the Molecular Physics"	Saving of the energy, electricity, and natural recourses	SPEM -8
TPEM -9	The physico-ecological potential of the theme "Heat Engines"	The energetics and ecology of the future	SPEM -9
TPEM -10	The ecological potential of the section "Thermodynamics"	The transports and environment	SPEM -10
TPEM -11	Electromagnetic phenomena and ecology	Physics problem of ecological aspect	SPEM -11
TPEM -12	The electromagnetic environmental pollution	Physics ecologically orientated experiment	SPEM -12
TPEM -13	The ecological potential of the theme "Electromagnetic Waves Scale"	The ecological consequences of the atmospheric disastrous phenomena	SPEM -13
TPEM -14	The ecological potential of the theme "Light Waves and Geometric Optics"	The biospheric physical phenomena and ecological situation	SPhEM -14
TPEM -15	The ecological element of the section "Atomic Physics'	The optical monitoring of the environment	SPEM -15
TPEM -16	The basics of the radiation pollution of the environment	The physical basics of dosimetry	SPEM -16
TPEM-17	The scientific, (biological, physical, ecological) picture of the world	Concerning the norms of keeping the environment clean and the requirements of the existing legislation	SPEM -17
		The measures of keeping the environment clean from physical pollution	SPEM -18

And how is the content of the physico-ecological module formed? During our research, the following criteria have been developed and are operating. According to the first criteria, it is important to introduce the educational content material in such a way that its lowest edge would correspond to the lowest threshold of the content of the material presented.

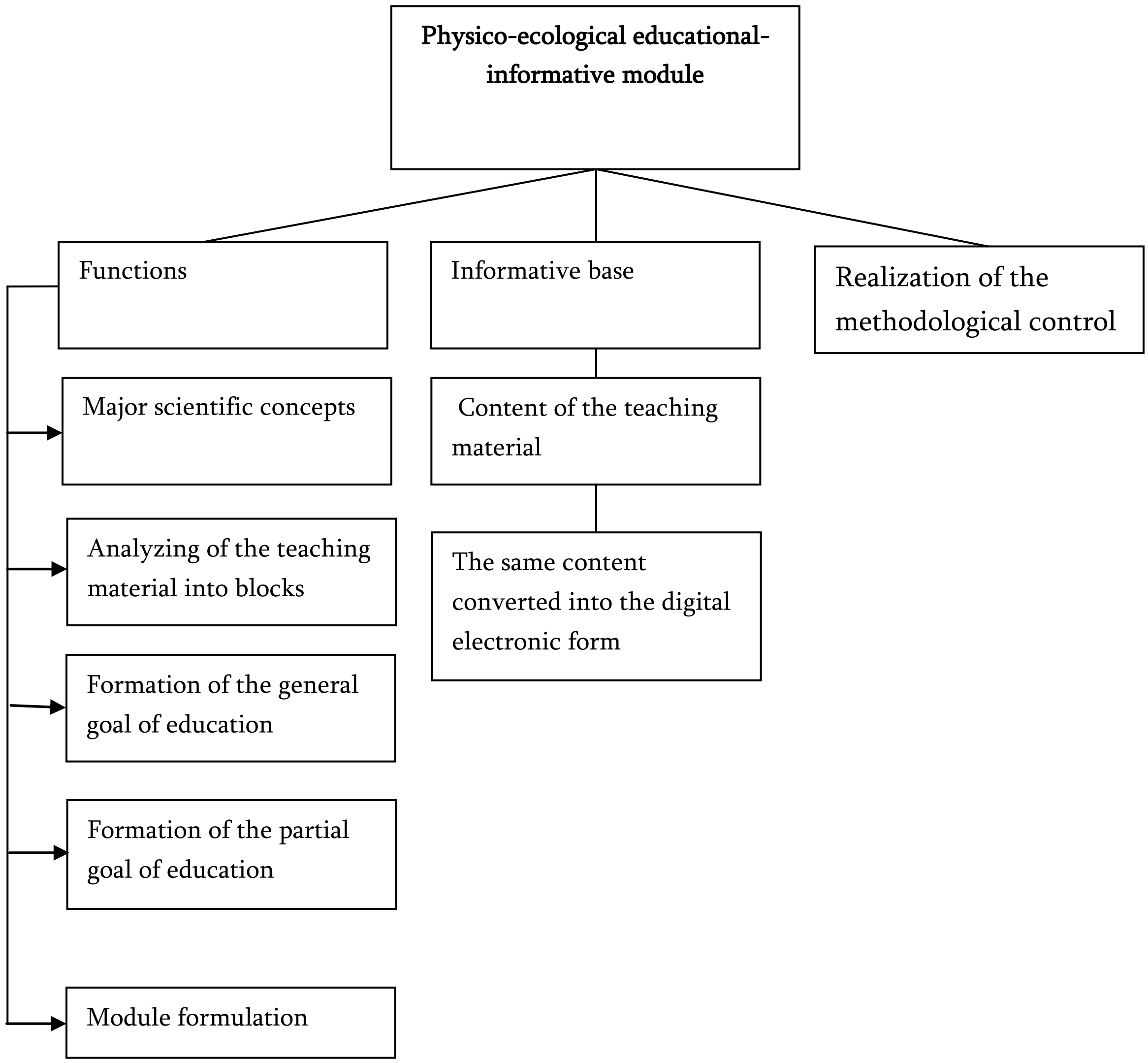
For the formation of the educational module, it is a very essential criteria to structure the learner's educational activity in the logical chain of the knowledge perception phases. In this case, it has a following form; to perceive $\rightarrow$ to comprehend $\rightarrow$ to think $\rightarrow$ to remember $\rightarrow$ to use $\rightarrow$ to generalize $\rightarrow$ to systemize.

Let's present the structure of the concrete informative teaching physico-ecological module. During the process of the module presentation, the following designations are used: EC-educational component, EP-educational problem, IDG-integral didactic goal, CDG-complex didactic goal, TPEM-thematic physic-ecological module, SPEM-special physic-ecological module.

During the organization of the modular education, if it simply refers to working by that technology, during the formation of the educational module, as a rule, numerous links are given both to the text-books, exercise-books and the existing educationally supported literature.

In case of a problem, the content- informative block of the presented educational module, covered issues, problems and accompanied academic information must be fully described and presented by the teacher.

Hence, the educational informative physico-ecological modules, considering structure, functions and their realization by methodological complex, could be presented by a following block-figure.



We present concrete samples of special and thematic physic-ecological modules

Special physico-ecological informative teaching module
"The physical sources of the environmental pollution"

Table 2

N/N	Educational information and educational task	Management of the educational activity
SPEM - 2	IDG: To perceive the following new concepts; "environment", "environmental pollution", "environmental physical pollution" EC 1. Environment and biosphere 2. Variants of environmental pollution 3. Noise pollution of environment 4. Thermal pollution of environment 5. Electromagnetic pollution of environment 6. Radiation pollution of environment 7. Measures for environmental pollution protection ES 1. Measuring of the noise level in the corridors of school and its surroundings. To make conclusions based on the measurement results (per a week) and to note the results on the school wall paper. 2. How to be sure about the global warming phenomenon? 3. How to measure the electrical fields created by the computer, photocopier and fridge. 4. What is a dosimeter? How does it work?	The task lasts 10-15 minutes The teacher presents in the form of a lecture. Its digital variant (if it has voice, it would be an advantage) would be desirable to provide to the learners. By the supervision of the physics teacher to carry out the measurements using the school materials if possible, otherwise to use a modern sound level meter, which can be obtained. To organize an excursion to Metsamore Nuclear Power Plant if possible.

Thematic physico-ecological module
“The Noise and Vibration Sources of the Environmental Pollution”

Table 3

N/N	Teaching information and teaching assignment	The management of the teaching activity
TPEM - 5	IDG To comprehend the following concepts and terms; “noise”, “vibration”, “noise level”, “noise pressure”, “sound pressure”, “sound intensivity”, “decibel”	The work lasts 10-15 minutes .

	EC 1. Environment and biosphere 2. How is noise produced? 3. What is noise? What is vibration? 4. Intensity of noise and its pressure 5. What does a noise and vibration pollution of environment mean? 6. Sound (noise) loudness level meter 7. Sound level meter, its structure and physical basics of its work	The physics teacher presents.
	ES 1. “Sound level meter, its structure, physical basics of its work” and “Measuring by sound level meter” and the realization of laboratory and practical work 2. Making measurements both in different places of school and in different places of the surroundings 3. To come to conclusions 4. To report about the results of the work to the members of the “Young ecologists” club	By the supervision of the physics teacher In collaboration with school scientific society

Different experiments, realized by us in different educational institutions, confirm that the integration of informative teaching of special and thematic physico-ecological modules in the teaching process of physics, not only increases the interest of learners to physics, but also contributes to the solution of the urgent issues of their ecological education and upbringing.

References :

[1] K. I. Atayan, Application of Integrative Modular Method of Physics in the Ecologized Teaching Process of Physics, Education and Sciences in Artsakh, **3-4**, 37 (2010).

[2] I. Zverev, Ecology in School Education. Moscow, Education (1980).

[3] B. Golubev, Dosimetry and Protection against Ionizing Radiation, A text-book for universities, edited by E. Stolyarova, Moscow, Energoatomizdat (1986).

[4] Y. Odum, Ecology, Volumes 1-2, Moscow, Mir, (1986)

[5] How should the environmental education be in 12-year secondary school? Round table, “Pedagogy”, **6,** 40 (2000).

[6] Y. Polezheva, Ecological Monitoring in the Teaching Process. Ecology and Life. **3,** 33 (2004).

[7] P. Yushcivengans, Theory and Practice of Modular Education, Kaunas (1989).

Increasing student understanding of Spectroscopy

and

Hertzsprung-Russell Diagrams

Arunava Roy

Division of Mathematics & Science
Young Harris College
Young Harris, GA USA.

(*Submitted 13-11-2012*)

Abstract

In this article the role of spectroscopy in understanding Hertzsprung-Russell diagrams is discussed. Intuitive methods of calculating equivalent and line widths of spectral lines are first presented. This is followed by a discussion justifying the use of line widths in temperature-luminosity diagrams.

1. Introduction

Among the more crucial topics in introductory astronomy are spectroscopy and Hertzsprung-Russell Diagrams (HRDs). Spectroscopy, the study of the light after passing through a prism or grating, reveals a wealth of information about the source including its temperature and composition; knowledge about some of these can only be obtained from spectroscopy. Introductory astronomy texts generally list spectroscopy in the beginning after a discussion of light, but its true application is revealed when studying properties of stars, including our sun. HRDs also come up when studying stellar properties, more specifically during a discussion of stellar life cycles. A HRD is a plot of, typically, luminosity and temperature of stars. Information about sizes and spectral classes of stars are some of the information obtained by analyzing HRDs. The aim of this article is twofold, firstly to introduce the concepts of equivalent width (EW) or the strength of a spectral line (including an easy measurement method using MS Excel) and line width (LW). Spectral width measurements are important because, for one, they allow us to determine the physical conditions in the atmospheres of stars. Moreover they let us quantify spectral lines and are rarely discussed in introductory texts. Simple algebraic calculations are employed to calculate the area under a curve and provide an intuitive way to calculate the EW. The article concludes with a discussion about the relationship between spectral features (i.e. EW & LW) and the HRD: that the HRD is actually a representation of the spectral features[1, 2, 3]. It emphasizes that the spectral features are actually dependent on the temperature and luminosity of a star-a critical concept that forms the physical basis for understanding HRDs.

2. Calculation of Line Widths

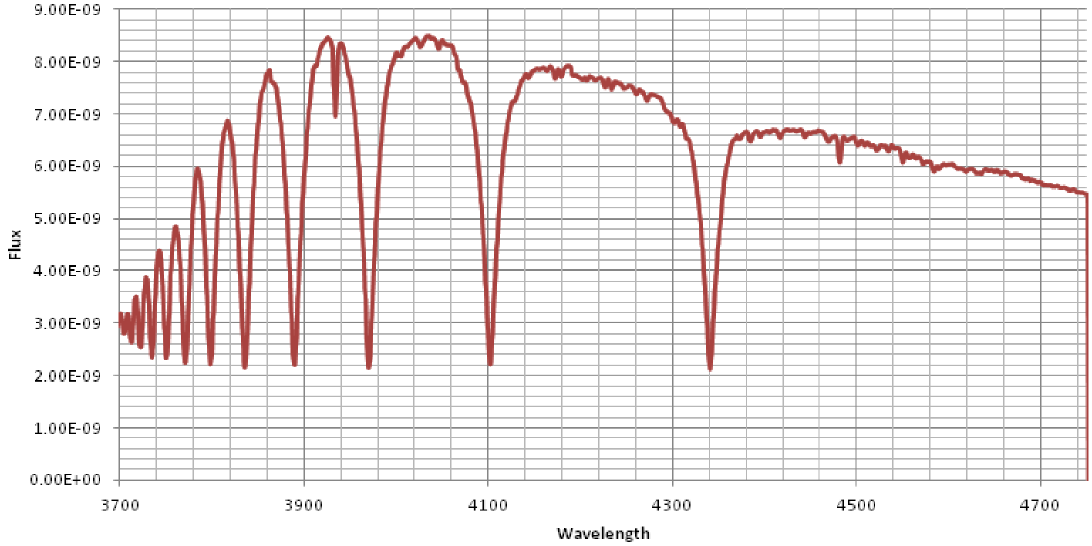


Figure 1: Sample spectra data obtained from STELIB and plotted using MS Excel showing the absorption features. Flux is expressed in solar luminosities per Angstrom.

Fig. 1 obtained using STELIB[4], shows a sample spectrum with flux plotted along the vertical axis and visible spectrum wavelengths along the horizontal. The dips in the plot are the absorption lines with the number of photons absorbed at a certain wavelength governing its shape.

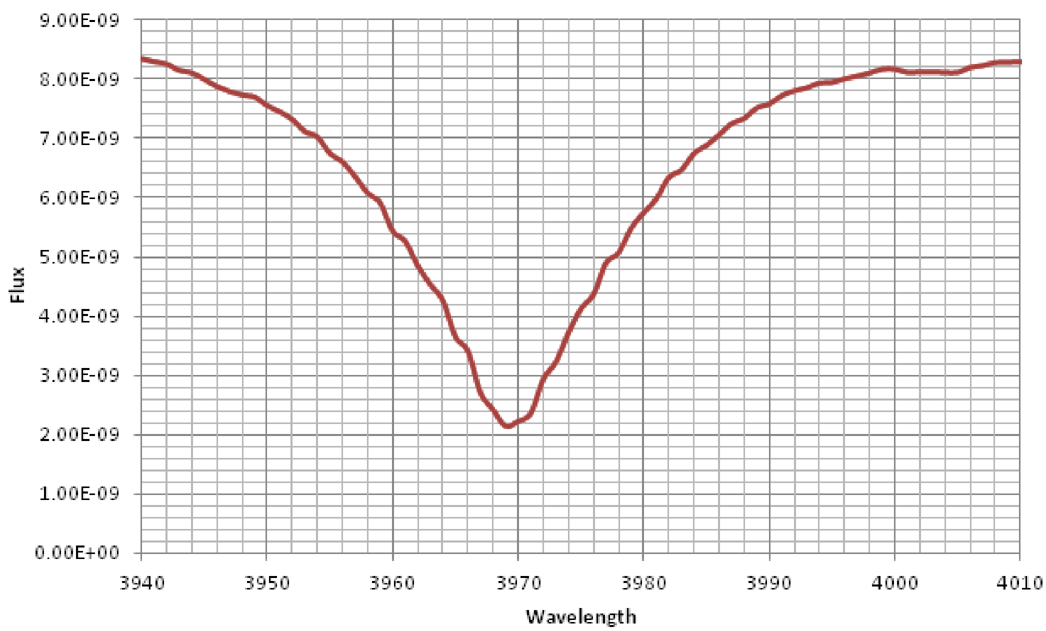


Figure 2: Same spectra as Fig. 1 but zoomed into the 3940-4010 A wavelength range. The equivalent width is calculated centered on 3969 A.

Fig. 2 is a snapshot of Fig. 1 in the wavelength range 3940A to 4010A. The choice of wavelengths is arbitrary and is just for the purpose of measuring the equivalent width of this particular absorption feature. Note that the peak wavelength occurs at about 3969A around which the EW would be measured. EW is defined as the width of the rectangle extending from the continuum to zero flux that has the same area as that enclosed by the spectra between the absorption and continuum lines. There are different ways of achieving this (apart from using calculus) and the methods mentioned herein offer a good approximation. More importantly, it gets the measurement idea across without employing any complicated mathematics. One way is to print the plot and count the number of boxes between the absorption and continuum line. A rectangle comprising of the same number of boxes could then be drawn around the peak wavelength extending from zero flux to the continuum, whose width would be the EW of the spectral line. A second method involves approximating the area in question by a triangle and calculating its area. Both of these methods are a visual way of calculating the EW and is therefore very appealing.

A third and more accurate method involves dividing the area under the curve into trapeziums[5]. The sum of the areas of the individual trapeziums is the area in question. This is illustrated in Fig. 3. For example, the area of the trapezium labeled 1 is calculated using the formula,

$$\text{Area of a trapezium} = \frac{1}{2} \text{ X sum of parallel sides X distance between the parallel sides.}$$

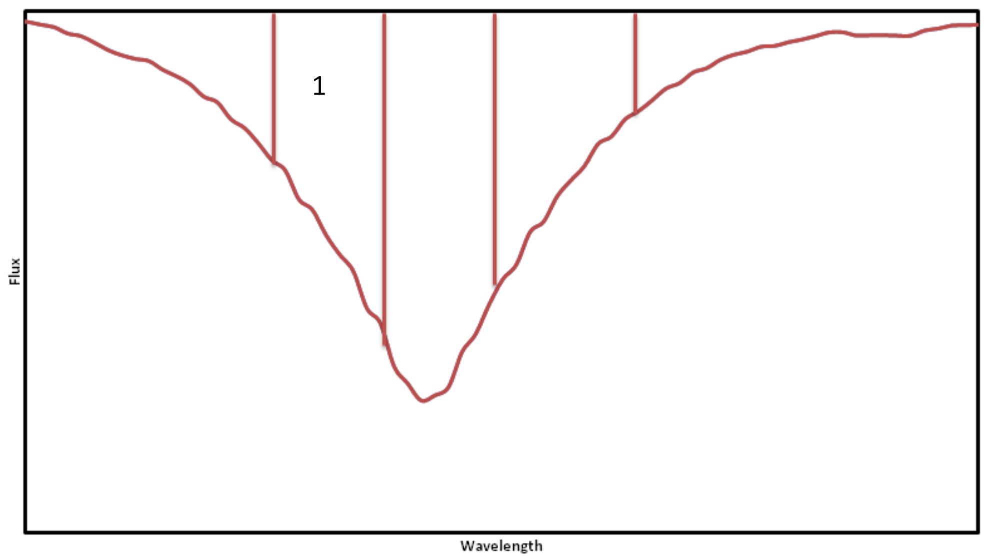


Figure 3: Calculating the area under a curve by dividing it into a number of trapeziums.

As an example, in Fig. 4, the area of the trapezium in the wavelength range 3950A-3952A is 0.22A. This process is automated using Excel until all trapeziums have been taken into account. Next, a rectangle centered on the peak wavelength is drawn to match the area calculated previously. Measuring the width of this rectangle is the EW of the line, as shown in Fig. 4. In this example, the

EW is approximately 15.5A around 3969A (as shown by the double headed arrow).

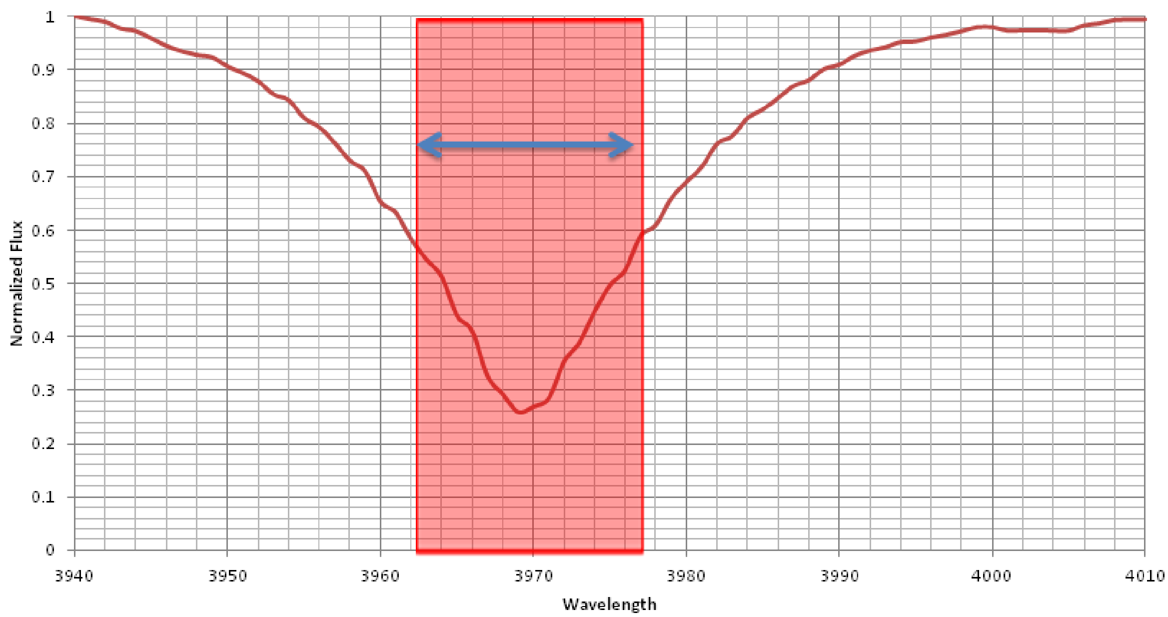


Figure 4: Shaded rectangle represents the equivalent width of the absorption line centered on 3969A. Note the vertical axis represents normalized flux.

Another way of quantitatively expressing the properties of a spectral line used by astronomers is LW. This is best outlined using an example. The solid and dashed spectral lines in Fig. 5 could have the same EW (i.e. same area under the each curve), but differ in the thickness around the central region. More precisely the width is measured at half the maximum vertical extent of the plot. This is a measure of the LW and it typically denoted by $\Delta\lambda$ in advanced literature. For a more technical definition of LW refer any astrophysics text[2].

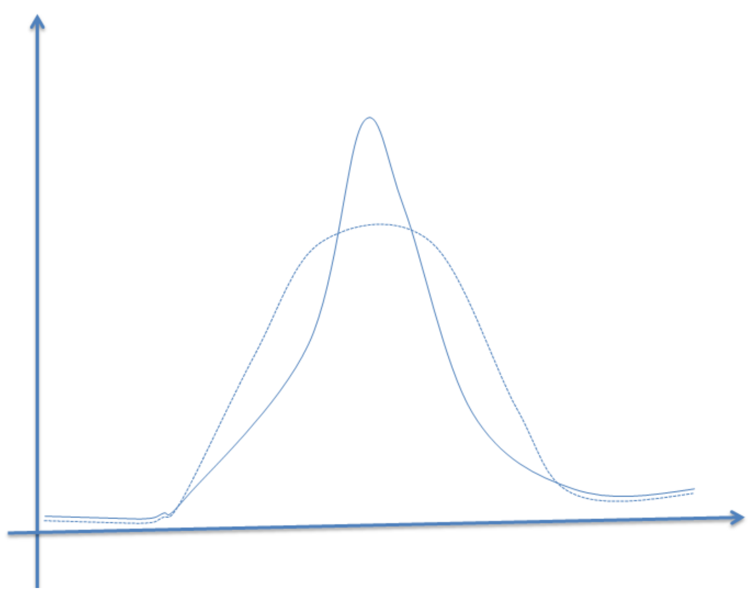


Figure 5: Solid and dashed spectral lines assumed to have the same equivalent width differ in line widths. Flux (wavelength) is along the vertical (horizontal) axis.

3. Relating Line Widths to HR Diagrams

We now proceed to relate spectral features (EW and LW) to our two dimensional HRD. The aim here is to reason how the spectral features are dependent on the temperature and luminosity. It starts with a simple question, how can we justify a plot of luminosity vs. temperature? The role of temperature in spectral line formation is well understood and appreciated by students. The line strength is dependent on the number of atoms making a certain atomic transition. The equations governing which atoms occupy a certain level and the number of ions formed are due to Boltzmann and Saha[2,3]. Together they provide an accurate description of the variation of line strength with temperature. The strength of the hydrogen Balmer lines, for example, peaks at 10^4K and decreases with either an increase or decrease in temperature. A lower temperature suggests not enough of the species available for making transitions and a rise in temperature favors ion formation, both of which are detrimental to formation of Balmer lines. Thus, the temperature dependence of the spectral features is established.

The other, not so much touted, factor that affects spectral line formation is the density of atoms, which other than being temperature dependent is also pressure dependent. It is also dependent on the composition of the star but we could ignore that for our purposes (by considering; for example, only Population I or II stars). Spectral lines vary in width with natural and Doppler broadening being the most important factors. Another type of broadening is known as pressure broadening. Spectral lines of two stars having the same temperature might not have the same width. A main sequence, giant or dwarf star of the same spectral type does indeed have different line widths. This is because a giant has a much less denser atmosphere (larger size) than a main sequence star. Low density in the photosphere implies less collisions with atoms (or ions) leading to less perturbations in the energy levels and hence less broadening of the spectral lines. As a result, spectral lines for the main sequence star would be broader than the corresponding giant.

4. Conclusions

Thus, other than the temperature dependence of spectral lines, the above discussion also suggests that stars with large surface area (and therefore brighter) have different widths than main sequence stars with the same temperature. Thus, the spectral features are also dependent on luminosity. Therefore, measurements of spectral features illustrate the role of spectroscopy in revealing the temperature and luminosity of stars, and their evolutionary status on an HRD.

References:

[1]. An Introduction to the Sun and Stars. Edited by Simon F. Green & Mark H. Jones. (Cambridge University Press).

[2]. Bradley W. Carroll & Dale A. Ostlie, An Introduction to Modern Astrophysics, 2nd Edition (Benjamin Cummings).

[3]. Stephen A. Gregory & Michael Zeilik, Introductory Astronomy and Astrophysics (Brooks Cole).

[4]. STELIB website, http://www.ast.obs-mip.fr/article181.html

[5]. A Simple Tool for Integration and Differentiation of Tabular Values in Microsoft Excel by Ole Anton Haugland, The Physics Teacher, December 2011, Volume 49, Issue 9, pp. 580.

The Young experiment as a teaching tool

N. Doumanis[1,3] and C. Trikalinos[2]

[1]European School of Brussels III, 135 BD du Triomphe 1050 Brussels.

[2]Athens University, Department of Philosophy and History of Science.
Panepistimioupoli Ilisia, 15771, Athens, GREECE
Corresponding author: Nikolaos Doumanis nikdoumanis@gmail.com

Abstract

The Young experiment is one of the leading experiments that clearly show the wave nature of light and at the same time, one can get quantitative results without using complicated calculations. As established by classroom experience, the measurements and the repeating of results leads students to very valuable conclusions concerning the understanding of nature, science and history.

KEYWORDS: Young Experiment, Optics, Interference

1. Introduction

The repeating of historical physics experiments and especially ~~of~~ those that led to the fundamental discoveries in physics plays a leading role in secondary teaching; in many cases even university courses [1, 2, 3]. Their power to deepen understanding might even be higher for those pupils or students that do not follow a scientific course.

Any learning procedure has as a final goal, the understanding of the material taught to the pupils/students. At this point many pedagogical approaches try to give answers. One possible tool, which in our opinion is powerful, is the repeating of selected historical experiments.

Each human society in history has developed a certain amount of knowledge and skills (as a mean in the total population), interacting continuously with scientific thought and tool invention. This means that the people we usually call thinkers, those that make a major contribution in philosophy, science etc. were confronted at each moment with the problems arising within their specific society and level of scientific evolution and tried to give the “best possible answer" in the gnoseological context of their era. Speaking especially for physics now, it is at that point that the concept of historical experiments emerges. We should first define what an historical experiment is. It can be conceived as an experiment that gives rise to a new theory, or solves (or contributes to the solution) of a controversy that existed at a specific time.

When dealing with such an experiment a pupil/student is invited to become restrained by the knowledge existing at the time; that is, to take into account only what was already known at the specific moment just before the historical experiment she/he will deal with. In this case the opportunity is opened for thinking about the conclusions she/he could come up with, if she/he was the one to discover for the very first time the corresponding results.

Based on the previous analysis, our approach is the following: we should first give a brief but clear understanding of the level of comprehension and scientific knowledge at the time of the experiment, then perform the experiment and analyze qualitatively as well as quantitatively the results obtained. After this, try to come to conclusions retracing more or less the thinking path of the scientists of that era. After teaching is completed, an evaluation of the results is of interest. It is, we believe, an issue as important as the first part of the procedure. Statistical results of selected questions can help us to understand the extent of comprehension and eventual misunderstandings of the pupils/students so as to improve the presentation (either experimental or theoretical). At that point we can summarize the above analysis as follows:

The cause of the significant importance of historical experiments is that, once properly introduced to the pupils/students, the latter may come to an understanding of: a) the connection of science and the acquired knowledge level of humanity with the socio-economical environment, b) the real essence of the problem studied and the role it plays on technology and the gnoseological level at which society is, c) the underlying physics and generally the importance of the expected results and d) the evolution of our knowledge and ideas in their recent form.

In such a process students do not gain knowledge, "from nowhere", but they understand the real adventure of the human mind in solving problems and acquiring knowledge. It seems that they will also gain deeper thinking capability and an extra interest in scientific problems. One can easily understand this point if one imagines the completely opposite situation. An experiment among those we defined as historical ones, if presented with few or without any explanations of the corresponding significant importance, means that students do not even see why they are making the effort to learn the experiment. In our opinion this is a total pedagogical disaster. It is thus always important to refer to such experiments, either in a normal course, or separately, to help students in the way described above.

One of the well known historical physics experiments is the double slit Young experiment. It is this experiment that we shall deal with in the present work.

2. The historical context

From the 17th century Newton's personality dominated in physics. His discoveries and ideas governed and any contestation of this was considered almost as a blasphemy. For more than 100 years, until about the end of the 18th century, his idea that light is constituted by particles remained dominant. At this point it is worthwhile giving the main lines of Newton's theory of light [4, 5]. He thought that light is constituted by small particles, very tiny ones that enter the human eye and cause what we call vision. Those tiny particles act like all other material bodies we know, such as stones or billiard balls. Of course a theory has to give an explanation of observed phenomena. Reflection of light is explained the same way as reflection of an elastic body against a wall. Refraction is a bit more complicated. When light particles come from the air into a transparent medium like water or glass for instance, with a velocity that makes a certain angle with the normal to the surface separating the two media, then entering the transparent medium their velocity suffers a change of angle (in fact it becomes smaller) compared to the previous one. In this case the gravitational attraction between light particles and the transparent medium increase the component of the velocity normal to the medium surface but leaves the parallel one unaffected. This explains why the entering angle of the velocity is less than the one with the surface medium but it also means that light should travel inside the new medium at a higher speed which, at that time was not verified. We know now that it is not the case and this is against Newton's theory. Still it is astonishing that Newton also gave an explanation, within the context of his theory, of why light should travel at constant velocity inside a medium: as the surrounding particles of the medium are exerting symmetrical gravitational forces to the light particles, they should give a resultant force equal to zero, as a mean value, at all times. He also gave an explanation of the difference of bending angles when light passes through a prism. He claimed that more massive light particles, those of red light, were deflected less than the lighter particles of the other colors, right down to the smallest of all those of blue color particles. At the end of the 18th century some doubts came up concerning the particle theory. In the context of a particle theory no interference phenomena were conceivable at that time, especially when the particles were considered as classic material particles. Therefore some scientists realized that in order to explain Grimaldi's (1618-1663) observations [6] concerning the diffraction of light, they had to use the wave theory of light. Grimaldi, in his two volume work, published posthumously in 1665, describes phenomena where the rectilinear motion of light beams, refraction and reflection cannot be applied to give an explanation of the experimental facts. He saw that when light passes through a hole the shadow produced is not the geometrical one. There is a clear and not unimportant region where light is present (something like a "brighter region into the geometrical shadow"). Also, near the edges he saw colors. He thus used the term diffraction and rejected the corpuscular theory. He thought of light rather as a fluid like water, where waves analogous with waves in water may propagate. These phenomena were known to Hooke and Newton and they used the term "inflection". Grimaldi's term diffraction is the one that finally survived. Others scientists followed [7] and the most important who investigated diffraction patterns was Fresnel (1788-1827) [6, 8, 9] who gave a full theoretical explanation of Grimaldi's observations.

Fresnel had a supporter, Arago (1786-1853) [6, 9]. They set up and performed a lot of experiments together. Some of them were very simple, but a number proved to be of great importance, because they convinced even skeptics or scientists opposed to the wave theory, that this theory in fact is true. Those experiments were influenced by Poisson (1781-1840) [6, 9], who was one of the greatest French scientists of the early 19th century.

Poisson claimed that - if light has a wave nature – and, in the path of a light beam which emerges from a point source we put an opaque circular disc perpendicular to the beam direction, the waves should come to every point of the circular obstacle with the same phase and thus give a luminous point at the centre of the obstacle's shadow! This (theoretical) result seemed to him completely absurd. But Fresnel with Arago performed this experiment and they observed that this luminous point in fact exists. From this moment, the theory of light's wave nature came to be generally accepted.

The most important conclusion of this event is that we should not ground our considerations and results only in the theory. No matter how obvious or elegant the theoretical predictions might appear, they should be tested through experiment. If this test gives satisfactory results relative to the theoretical predictions the corresponding theory may be accepted and new experiments can be set up. But if the theory does not explain the experimental facts, the theory might be revised or be changed completely. In the case of physics all these possibilities are open.

The wave nature of light can be tested in a different way, by using the interference phenomenon. The most famous experiment is the one performed by Young (1773-1829) [10, 11] who observed light interference when light passes through two narrow slits. Young understood that it is impossible to observe interference phenomena with light coming from two independent sources and for this reason in 1807 performed the following experiment: in a dark room he let sunlight pass through a very narrow hole that he produced with the aid of a very fine needle, thus producing a diverging sunlight cone. In the center of this cone he put a piece of paper about one millimeter in size, thus dividing this beam into two parts. Afterwards he put a screen in the direction of the split~~ed~~ beam and saw interference fringes on it that were symmetrically placed. The central fringe was white. At the edges fringes were colored. The fringes thus produced coincided with the corresponding fringes that Fresnel saw. When Young moved the piece of paper towards the edge of the initial beam the interference fringes disappeared. Young performed many experiments in order to convince himself that this phenomenon appears when the initial beam is divided into two beams.

The simplicity as well as the persuasiveness of Young's experiment played a very important role in the confirmation of Fresnel's work concerning the wave theory of light.

The most important advantage of Young's experiment is that, it can furnish quantitative results with the use of simple mathematics. Fresnel's diffraction also can give quantitave results, but in this case rather complicated mathematical calculations are needed. The Young interference fringes can be described by a simple theory and be used to measure wavelengths of light with an accuracy of about 1%.

3. The experiment

Repeating this experiment is nowadays a standard procedure which is used in school classes, usually at the final level of the secondary, as at this level students are more mature and able to follow, at least in principle, the mainlines of the experiment as well as the scientific reasoning.

To prepare one of these experimental presentations we performed the Young experiment as follows (students were not present at this stage). The experiment was performed in the laboratory of the European School Brussels III. We will describe the teacher's preparation and then present the part which, at this level, will be presented to the students:

We used as light source a 5 beam He-Ne LASER (figure 1), of 2 mW power and a wavelength of 632.8 nm, as provided by the factory. Actually we take this wavelength as being error free. The same thing holds for the two slits (figure 2) which are at distance of 0,6 mm, as provided by the factory.

Figure 1. The He-Ne LASER we used

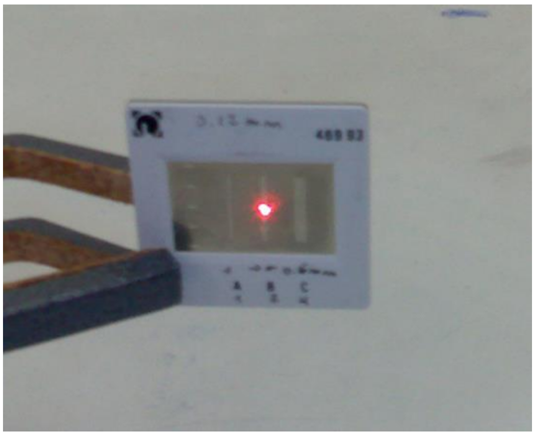


Figure 2. The LASER beam passing through the double slit.

The LASER beam was at a distance of 20 cm from the double slit (figure 3) and the distance of the double slit to the screen (actually a wall) was of $(250,5 \pm 0,5)$ cm. This error is an estimated error.

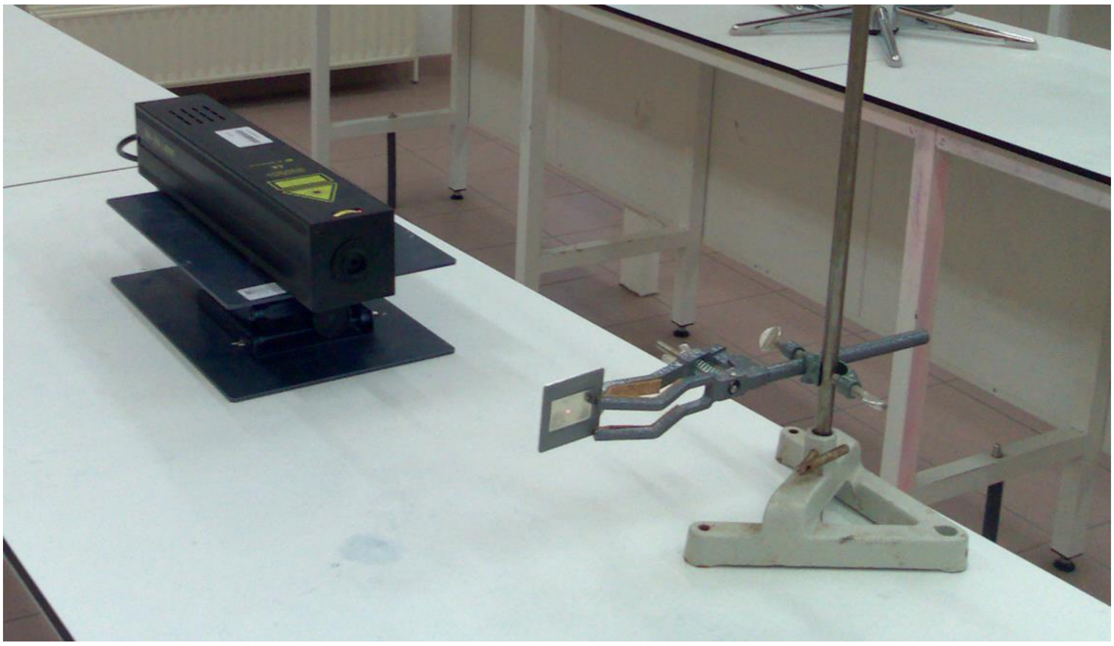


Figure 3. The LASER together with the double slit

4. Results and calculations

The goal of the experiment is: a) the qualitative presentation of the results of the Young experiment and b) the corroboration of the corresponding formula, which we will first derive.

Suppose that light, from a monochromatic source comes to a double slit, as shown in figure 4. We denote by a the spacing between the two slits, D is the distance between the double slit plane and the screen. The two slits are denoted by L, S respectively. The line MO is perpendicular to a at the midlepoint M and to the screen at the point O. OP is the distance of the constructive interference fringe of order k, which we denote by x. We draw two lines LP_1 and SP_2 normal to the screen.

Obviously from the two right angled triangles LPP_1 and SPP_2 using Pythagoras's theorem we have the following relations:

$$(SP)^2 = D^2 + \left(x + \frac{a}{2}\right)^2 \qquad (1)$$

And:

$$(LP)^2 = D^2 + \left(x - \frac{a}{2}\right)^2 \qquad 2)$$

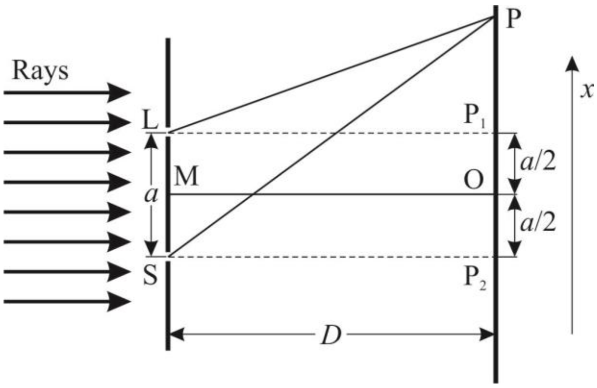


Figure 4. Schematic diagram used in the Young formula demonstration

Subtracting (2) from (1) and performing trivial calculation we get:

$$\left[(SP)-(LP)\right]\left[(SP)+(LP)\right]=2xa \qquad (3)$$

The second term in brackets, in the first member of the equation, is about 2D as $D >> a$, and the distance x of any of the fringes is very small compared to D. So without any significant error we may set:

$$(SP)+(LP)\cong 2D \qquad (4)$$

It is well known that when light comes from coherent sources a point like P is a constructive interference point of order k if the path difference is k times the wavelength λ. So it must be:

$$(SP)-(LP)=k\cdot\lambda \qquad (5)$$

Substituting (4) and (5) into equation (3) we get the well known Young formula:

$$\lambda=\frac{a\cdot x}{k\cdot D}$$

where, to avoid any confusion, λ is the wavelength of the light, a is the distance between the two slits, x is the distance of the luminous fringe of order k from the central one, and finally D is the distance between the double slit plane and the screen (actually the wall). It is obvious that the measured quantity is x.

With regard to the results, we concentrated on the fringes near the central luminous fringe. We observed 4 luminous fringes at each side of the central one (figure 4). Thus the total number of luminous fringes was 9. The distance between the two edge fringes (right and left side) was $(24,0\pm 0,2)$ mm. This error is also an estimated error.

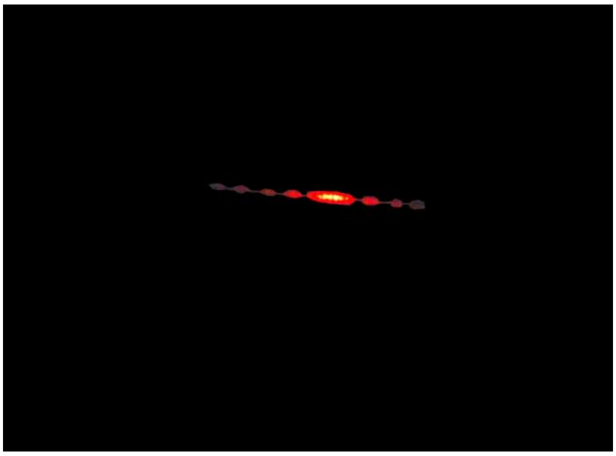


Figure 5. The observed luminous fringes on the wall.

The width of all 9 luminous fringes seemed to be the same, as well as the width of the dark fringes between them. The width of the dark fringes, which is easier to measure, was found to be the same for all of them, as $(1,0\pm 0,2)$ mm. To find the distance, according to our measurements, between the central luminous fringe and one luminous fringe of higher order, we took into account that the measured total width of the eight dark fringes was $(8,0\pm 0,2)$ mm, which means that the $(16,0\pm 0,3)$ mm, until the total of $(24,0\pm 0,2)$ mm measured, corresponds to the total width of the luminous fringes. As the width of all luminous fringes seemed to be the same (no better measurement was possible), we concluded, dividing by 9 (which is the total number of luminous fringes), that the width of each luminous fringe is 1,8 mm.

Thus to do our control on the reliability of the performed experiment, we should calculate the distances *x* of the different order fringes and compare them with the experimental values.

As already mentioned, the LASER wavelength and the slit distance are known and considered as error free – values provided by the factory.

In the first table below (TABLE 1) the experimental results are presented together with the theoretical calculations and the corresponding errors. The errors presented in this table result from the measurement precision and the theory of error calculation.

Table 1 Distance *x* [mm] until fourth order

k	Theoretical results	Experimental results
1	2.64±0,02	2.8±0.2
2	5.28±0.05	5.6±0.3
3	7.93±0.07	8’4±0.4
4	10.57±0.10	11.1±0.4

Comparing the calculated, (according to the theory) and the experimental results, we see that the theoretical results are in the range of the experimental values; we can therefore consider the experiment as being of good reliability.

5. Student’s presentation

We think that our conclusions would be incomplete if we did not include comment on the experience we have had, until now, of teaching Young’s experiment in class and explaining how this publication can help achieve better teaching in class.

It is obvious that in the classroom the overall problem that has to be solved is that students should, in the end, come to an understanding of light as a wave through the results of Young’s experiment. In our practical experience we have seen that without the historical presentation students did not understand why this experiment is so important and how it, in fact, contributed to a deeper understanding of light as a wave.

With respect to the experiment itself and the corresponding explanation of the calculations and the approximations used we did not have a major problem. As the source is practically monochromatic (He-Ne LASER), and the two slits very close (values already given above) the students did not have any problem understanding the calculations nor the approximations used because they observed how much bigger the distance LASER – screen was compared to the distance of the two slits between them.

When they try to measure the distance of the light fringes from the central one, and also measure the distance of the LASER from the screen, they begin to understand that this is not a trivial task, but requires a series of measurements which should be done with caution and repeated several times. At this moment it is a good idea to reintroduce the notion of the mean value of several measurements.

In the secondary program, in most cases, no analytic theory of experimental errors is included; in the presentation which was dedicated to the students we made only qualitative considerations concerning the experimental errors.

We should thus present the results as they are in table 2:

Table 2. Distance x [mm] until fourth order

k	Theoretical results	Experimental results
1	2.6	2.8±0.2
2	5.3	5.6±0.3
3	7.9	8.4±0.4
4	10.6	11.1±0.4

This means that the students should become, once again, aware of the fact that experimental error always exists and they should be invited to think where experimental errors come from in the specific case of Young's double slit experiment.

The next point, which is of major importance, is to make students think why Young's experiment is of such importance in Physics and thus one of the experiments very often included in secondary physics courses (This is in fact already explained in our introduction above). Students should understand that it was a simple experiment which proved the wave nature of light, giving at the same time the possibility of simple calculations which ensure that theory is describing reality accurately. To this end, we can make them calculate with measured distances from the central fringe, the LASER wavelength. Using the above mentioned values we find: $\lambda = 670$ nm. This has a relative difference of about 6% compared to the value given by the manufacturer. Thus the students understand that this method can also give quantitave results.

6. Students reaction and comprehension

The main goals of the Young experiment that we repeated in class are: i) students should get aware of the fact that light is a wave, that is that the experimental facts at the period that Young's experiment was performed were in favor of the wave nature of light, ii) students should also understand the analysis of the experimental procedure in detail, the calculations involved to obtain the results and the approximations and how these approximations can be justified, iii) the historical context, that is presentation in a more detailed way of the theories that existed at Young's time and why his experiment played a decisive role in favor of the wave nature of light (this point completes point (i) above).

To substantiate, by using a group of questions, the percentage of our goals which were achieved, we prepared a questionnaire. This questionnaire is as follows:

1. Did this experiment help you to understand why light presents the characteristics of a wave?
   a. Very much, b. Enough, c. A little, d. Not at all.
2. Did you understand the historical context and the controversy about light's nature at that time?
   a. Very much, b. Enough, c. A little, d. Not at all.
3. Did you understand the calculations that we used
   a. Very much, b. Enough, c. A little, d. Not at all
4. When performing the experiment we used a monochromatic (LASER) beam. Did you understand the difference with the procedure used by Young?
   a. Very much, b. Enough, c. A little, d. Not at all.

5. In your opinion which is the most important approximation that we used in our calculations?
6. Which are in your opinion the most important sources of errors in this experiment?
7. Why did we refer to Huygens's principle while discussing this experiment?

Taking into account the students responses to the questions we concluded that: A percentage close to 85% of the students understood very well, or well enough, the wave nature of light, whilst about 60% had a good understanding of the historical context and the controversies concerning the nature of light at that time. It is remarkable that only a few of them came to a complete understanding (11%) of the difference between the experiment performed by Young himself and the one we did in class using a LASER beam of practically one wavelength.

The students also had difficulties in understanding the main approximations used, the origins of experimental errors, and why we referred to the Huygens principle in our explanations.

In order to compare the results and deepen our understanding on students difficulties, we used the same questionnaire to another group of students this year (2013). Taking into account the students responses we concluded that: A percentage of 93% of the students understood very well, or well enough, the wave nature of light, whilst almost 70% understood the historical context and the controversies concerning the nature of light, at that time. It is remarkable that in this second group almost 70% of the students understood completely or well enough the difference between the class experiment and the one performed by Young himself.

In this second group the difficulties in understanding the main approximations used, the origins of the experimental errors, and why we referred to the Huygens principle, are restricted to only about 25% of the students.

If we compare the results obtained by the two groups of students, it is clear that the second group understood better all the points that constitute our goals. This fact is worth of an analysis of the reasons that led to a better performance the second group.

We presented first a brief presentation of the corpuscle theory of Newton and Grimaldi's point of view that light should be a wave. We did not "solve" at that moment the controversy between the two theories, we proceeded instead to the experiment which we presented with all the necessary details, concerning the experimental procedure itself as well as calculations and approximations used.

Only after this presentation we explained why Young's experiment leads to the conclusion that light is a wave and thus solve the controversy in favor of the wave theory.

It seems that the results obtained this way are much better than the previous ones and thus the procedure to be used is the one just described. Of course other teachers that might use our work as a guideline should check carefully the steps to be used and confirm or modify the described procedure.

7. Conclusions

If wave optics is included within the school syllabus, this experiment is among the most important. The goal of this experiment (and the analysis made of the results) is to make the students understand that the experiment (the one of Young in this case), gave clear evidence that light is a wave within a given historical context.

A deeper understanding of present day knowledge will come later on, perhaps in a University course, when the students find out that a similar double slit experiment reveals a wave nature for electrons. If they already know and understand Young's experiment they will find it easier to understand how this question of the duality wave-corpuscle is resolved through the statistical nature of the probability of finding a massive or non massive particle at a given point of space and at given moment of time.

Although the issue of the wave nature of material particles is rather a matter of a university course, a first contact with this subject can be achieved in secondary education. This can be done by experiment, if the necessary equipment exists, or by a computer simulation that exists on several internet sites which can be used to make clear that a wavelike pattern can also be obtained with material particles, such as electrons. Through combined simulations we also have the possibility of comparing the two cases [12, 13].

One of the most difficult points for the students turns out to be the historical significance of this experiment. Therefore the class teacher has to make clear that once the general rules of constructive wave interference are applied in this case, (they should obviously have already been taught in class), we find results that approach, to a very high accuracy , the observed data. This means that in the beginning we just suppose that light is constituted by waves and then we get results that justify this initial hypothesis.

One last but very important point is to explain, perhaps briefly but in a clear way, how Young managed to do such an experiment with the poor means that he had at that time. Our introduction above can guide the class teacher on that, but it is obvious from the results of our questionnaire that any subject teacher has to explain, as rigorously as possible, the monochromatic nature of the beam used in today's classroom. This description should not be dealt with first. It is more instructive to present how Young managed to do his experiment once the experiment with the LASER is already done and rigorously explained in class.

As already mentioned, the controversy of wave-corpuscle is one which dominated in the past. Nowadays it still very important, especially for people that are not engaged with natural sciences, and especially physics, as it may be their only contact with this matter.

For all students, even those following human sciences, the methodological approach to Physics problems, as expostulated above, is of great importance, as is an understanding of the pioneers of science and today's researchers. Given this, we should spend more time in class and find better approaches for the presentation of approximations and experimental errors. A first try is the one we made with the second group where we managed to achieve better results. This effort should continue by other professors. If the reader wants to do some further reading and research for himself, he is invited to use also some further bibliography [14].

Acknowledgments

We have to thank Naim Bajraktari, who is responsible for the Physics Laboratory of the European School Brussels III, who took care in providing the necessary material and helped throughout the experimental process.

We also have to thank the English Section professor of the European School "Brussels III" Mr. Peter Cox who made numerous remarks on the manuscript. His remarks improved substantially the final version of this publication.

References

[1]. Mathews M R, 1994 *Science Teaching – The role of History and philosophy of science,* New York, London, Routledge

[2]. Max Born and Emil Wolf – Principles of Optics 1999 Cambridge University Press

[3]. OS Heavens and RN Ditchburn "Insight into Optics" John Wiley and sons Chichester

[4]. Cajori F, 1899 *A History of Physics in its Elementary Branches, including the evolution of physical laboratories*. New York, MacMillan Company

[5]. Isaac Newton Optics, 1704, fourth edition 1730, reprinted by Dover publications 1952

[6]. Segre E, 1984 *from Falling Bodies to Radio Waves: Classical Physicists and their Discoveries.* New York: W.H. Freeman,

[7]. Maraldi G. F. "Diverses Experiences d'Optique" Memoire de L' Academie Royale des Sciences, 1723

[8]. Augustin Fresnel Œuvres Complètes 1 Paris Imprimerie Imperiale 1868

[9]. Shamos M H, 1987 *Great experiments in physics: firsthand accounts from Galileo to Einstein*, New York, Dover Publications

[10]. Walter Scheider "The Physics Teacher" 24 217-219. 1986

[11]. Young T, 1804 *Experimental Demonstration of the General Law of the Interference of* (Reprinted in [5] pp 96-101)

[12]. Internet: Physicsclips 4UNSW School of Physics Sydney Australia Young experiment – water waves and light waves (http://www.animations.physics.edu.au/light/nature-of-light/index.html)

[13]. Internet: Double slit experiment (electrons) – Hitachi (http://www.youtube.com//watch?=ZJ-0PB)

[14]. Thomas Young Experiment, bibliography (http://thomasyandlight.edublogs.org/category/bibliography)