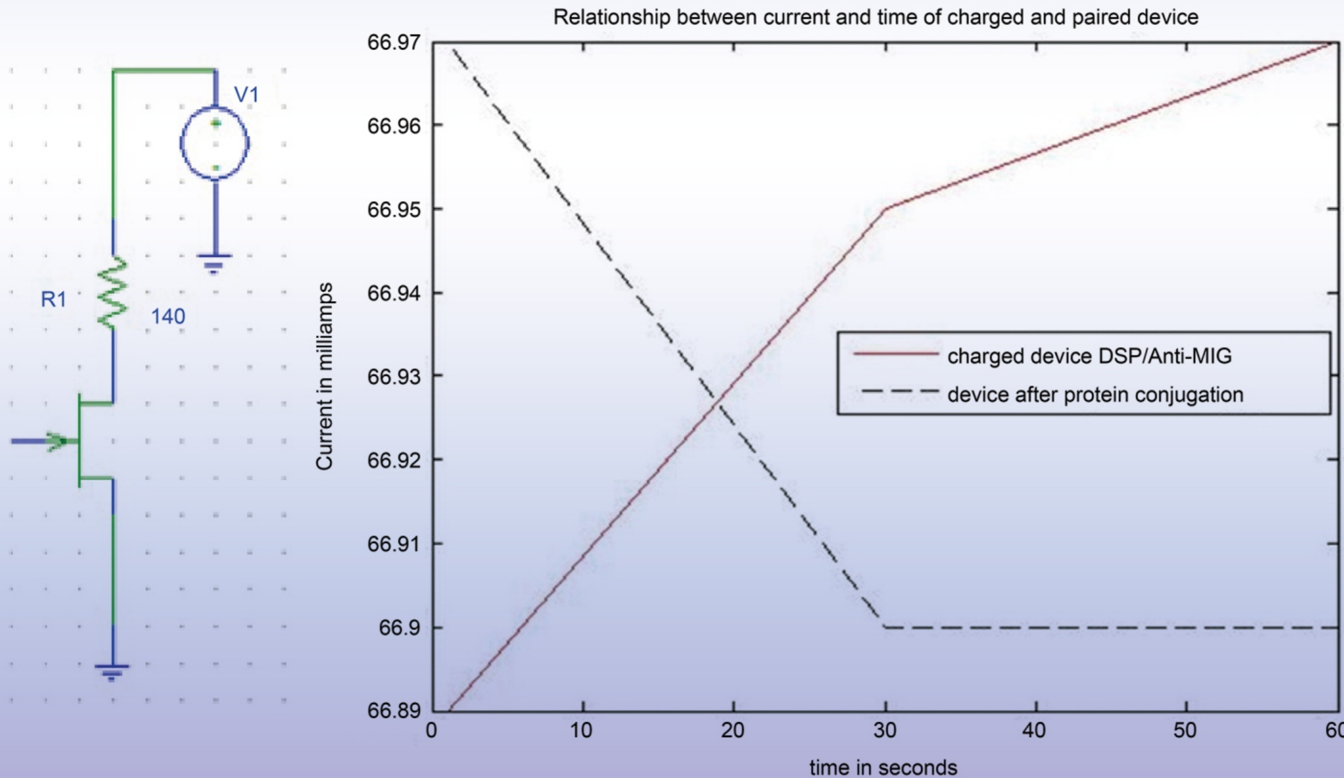
ISSN: 2153-1196 Volume 7, Number 13, September 2016

Scientific Research Publishing

# Journal of Modern Physics




ISSN: 2153-1196

www.scirp.org/journal/jmp

# Journal Editorial Board

ISSN: 2153-1196 (Print) ISSN: 2153-120X (Online)
http://www.scirp.org/journal/jmp

Journal of Modern Physics, 2016, 7, 1617-1844
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Table of Contents

**Volume 7 Number 13 September 2016**

Journal of Modern Physics, 2016, 7, 1617-1844
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196



The figure on the front cover is from the article published in Journal of Modern Physics, 2016, Vol. 7, No. 13, pp. 1712-1724 by Hasina F. Huq, *et al.*

# Journal of Modern Physics (JMP)
# Journal Information

## SUBSCRIPTIONS

The *Journal of Modern Physics* (Online at Scientific Research Publishing, www.SciRP.org) is published monthly by Scientific Research Publishing, Inc., USA.

**Subscription rates:**
Print: $89 per issue.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

## SERVICES

**Advertisements**
Advertisement Sales Department, E-mail: service@scirp.org

**Reprints (minimum quantity 100 copies)**
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: sub@scirp.org

## COPYRIGHT



## PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: jmp@scirp.org

Journal of Modern Physics, 2016, 7, 1617-1626
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Gauge Invariance of Gravitodynamical Potentials in the Jefimenko's Generalized Theory of Gravitation

**Augusto Espinoza, Andrew Chubykalo, David Perez Carlos**

Unidad Académica de Física, Universidad Autónoma de Zacatecas, Zacatecas, México
Email: achubykalo@yahoo.com.mx

**How to cite this paper:** Espinoza, A., Chubykalo, A. and Carlos, D.P. (2016) Gauge Invariance of Gravitodynamical Potentials in the Jefimenko's Generalized Theory of Gravitation. *Journal of Modern Physics*, **7**, 1617-1626.
http://dx.doi.org/10.4236/jmp.2016.713146

**Received:** July 5, 2016
**Accepted:** September 3, 2016
**Published:** September 6, 2016



## Abstract

In the Jefimenko's generalized theory of gravitation, it is proposed the existence of certain potentials to help us to calculate the gravitational and cogravitational fields, such potentials are also presumed non-invariant under certain gauge transformations. In return, we propose that there is a way to perform the calculation of certain potentials that can be derived without using some kind of gauge transformation, and to achieve this we apply the Helmholtz's theorem. This procedure leads to the conclusion that both gravitational and cogravitational fields propagate simultaneously in a delayed and in an instant manner. On the other hand, it is also concluded that these potentials thus obtained can be real physical quantities, unlike potentials obtained by Jefimenko, which are only used as a mathematical tool for calculating gravitational and cogravitational fields.

## Keywords

Gravitational Potentials, Cogravitation, Helmholtz's Theorem

## 1. Introduction

Jefimenko's generalization of Newton's gravitational theory [1] [2] is based to a large extent on the assumption that there exists a *second* gravitational field (which he has named the cogravitational, or Heaviside's, field). Note that there are several publications in which it is suggested that a second field can be involved in gravitational interactions (see [3] and, e.g., [4] and references there). The *first* such publication was by Oliver Heaviside [3], unfortunately his article appears to have been generally ignored (see, e.g. [4] pp. 103-104). The overriding reason why Heaviside's work did not attract the attention was that his single article on gravitation was eventually completely eclipsed by Einstein's brilliant and spectacularly successful general relativity theory. It is noteworthy, however, that Newton's gravitational theory generalized to time-dependent systems yields several results which heretofore are believed the exclusive consequence of the general relativity theory. Jefimenko discusses this very important circumstance in [1] [2]. It is interesting to note that Einstein, four years before he published his general relativity theory, published an article on the possibility of a gravitational analogue of electromagnetic induction. Also this article was practically unknown, possibly because it was published in a rather inappropriate journal whose title (translated from the German) was: "Quarterly Journal for Forensic Medicine and Public Sanitation;" (see [5]).

Newton's theory does not include inductive phenomena, but a relativistic theory of gravitation should include them. Indeed, under the relativistic mass-energy equivalence, not only the mass is a source of gravitational field but any kind of energy also is. Therefore, a body creates gravitational field not only by mass but also by their kinetic energy, *i.e.* by their movement. And this, ultimately, is what it means induction: the production of forces by moving bodies [6].

In the general relativity theory, Einstein predicted the existence of gravitational induction phenomena, such phenomena are appointed by Einstein as gravitomagnetism. It can be showed that Jefimenko equations are also derived from linearized Einstein equations (see, for example, pp 47 and 48 in [6]).

Since in 2004, NASA has orbited the "Gravity Probe B", whose purpose was to prove the existence of gravitomagnetism, (see [7] [8]).

In order to describe the time-dependent gravitational systems, the Jefimenko's generalized theory of gravitation is based on postulating of retarded expressions for the accustomed gravitational field $\boldsymbol{g}$ and the Heaviside's or cogravitational field $\boldsymbol{K}$ (Heaviside [3] was the first who supposed the existence of this field making an analogy between gravitational and electromagnetic fields), where the field $\boldsymbol{g}$ acts to and arises from motionless as well as moving masses, and the field $\boldsymbol{K}$ acts to and arises from exclusively moving masses. Let us from this point call this theory "gravitodynamics", and the complex of the fields $\boldsymbol{g}$ and $\boldsymbol{K}$ call the "gravitodynamical field". Jefimenko taking into account mentioned retarded expressions for $\boldsymbol{g}$ and $\boldsymbol{K}$, obtains the system of equations for which these expressions are solutions. These equations are analogous to Maxwell's equations. In principle, this method was proposed to eliminate the possibility of instantaneous solutions from the discussion. There against in this work we are going to postulate the system of differential equations rather than solutions for the gravitational dynamics, and we will obtain both retarded and instantaneous solutions for the fields $\boldsymbol{g}$ and $\boldsymbol{K}$. The gravitational field $\boldsymbol{g}$ behaves analogously to the electric field in the Maxwell's electromagnetic theory, and cogravitational field $\boldsymbol{K}$ analogously to the magnetic field.

First of all, we write the equations describing time-dependent gravitational systems [1] (see p. 120) and [2]

$$\nabla \cdot \boldsymbol{g} = -4\pi G \varrho \tag{1}$$

$$\nabla \cdot \boldsymbol{K} = 0, \tag{2}$$

$$\nabla \times \boldsymbol{g} = -\frac{\partial \boldsymbol{K}}{\partial t}, \tag{3}$$

$$\nabla \times \boldsymbol{K} = \frac{1}{c^2}\frac{\partial \boldsymbol{g}}{\partial t} - \frac{4\pi G}{c^2}\boldsymbol{J}, \tag{4}$$

where $c$ is the velocity of propagation of the fields, which is supposed equal to the velocity of light for the retarded component, $G$ is the constant of gravitation, $\varrho$ is the density of mass and $\boldsymbol{J} = \varrho \boldsymbol{v}$ is the density of mass current.

There are some differences between Maxwell's equations of electrodynamics and the Jefimenko's equations of gravitation, *i.e.* the analogy is not perfect. For example, we have two kinds of electric charges, positives and negatives, which repel each other if the charges are equal and attract each other if they are different, whereas while we have only one type of mass, and if we have a system of two masses in repose, they always attract each other. While the electric field is directed from positive charges generating this field and is directed to the negative charges, the gravitational field is always directed to the masses by which is created. Another difference is that the magnetic field is always right-handed relative to the electric current by which is created, while the cogravitational field is always left-handed relative to the mass current by which is created.

In the analogy between electrodynamics and the so-called gravitodynamics, following the Jefimenko's book [1] we can resume the correspondence between electromagnetic and gravitodynamic symbols and constants in the following Table 1.

## 2. The Gravitodynamical Potentials

Here, we introduce as is made in electrodynamics, the gravitodynamical potentials. If

**Table 1.** Corresponding electromagnetic and gravitodynamic symbols and constants.

| Electromagnetic | Gravitational |
|---|---|
| $q$ (charge) | $m$ (mass) |
| $\varrho$ (volume charge density) | $\varrho$ (volume mass density) |
| $\sigma$ (surface charge density) | $\sigma$ (surface mass density) |
| $\lambda_e$ (line charge density) | $\lambda_m$ (line mass density) |
| $\varphi$ (electric's scalar potential) | $\tau$ (mass's scalar potenctial) |
| $\boldsymbol{A}$ (magnetic vector potential) | $\boldsymbol{\Gamma}$ (cogravitational vector potential) |
| $\boldsymbol{J}_e$ (convection current density) | $\boldsymbol{J}$ (mass-current density) |
| $I_e$ (electric current) | $I$ (mass current) |
| $\boldsymbol{m}$ (magnetic dipole moment) | $\boldsymbol{d}$ (cogravitational moment) |
| $\boldsymbol{E}$ (electric field) | $\boldsymbol{g}$ (gravitational field) |
| $\boldsymbol{B}$ (magnetic field) | $\boldsymbol{K}$ (cogravitational field) |
| $\varepsilon_0$ (permittivity of space) | $-1/4\pi G$ |
| $\mu_o$ (permeability of space) | $-4\pi G/c^2$ |
| $-1/4\pi\varepsilon_0 = -\mu_o c^2/4\pi$ | $G$ (gravitational constant) |

the cogravitational field $\boldsymbol{K}$ satisfies Equation (2), we can always write it as the curl of some other vector quantity $\boldsymbol{\Gamma}$,

$$\boldsymbol{K} = \nabla \times \boldsymbol{\Gamma}, \tag{5}$$

where $\boldsymbol{\Gamma}$ is the gravitodynamical vector potential. Substituting Equation (5) in (3), we obtain

$$\nabla \times \left( \boldsymbol{g} + \frac{\partial \boldsymbol{\Gamma}}{\partial t} \right) = 0. \tag{6}$$

The quantity within the parentheses can be written as the gradient of a gravitodynamical scalar potential $\tau$:

$$\boldsymbol{g} + \frac{\partial \boldsymbol{\Gamma}}{\partial t} = -\nabla \tau, \tag{7}$$

therefore,

$$\boldsymbol{g} = -\nabla \tau - \frac{\partial \boldsymbol{\Gamma}}{\partial t}. \tag{8}$$

Substituting the expressions (5) and (8) for the fields $\boldsymbol{g}$ and $\boldsymbol{K}$, in the inhomogeneous Equations (1) and (4), we obtain

$$\nabla^2 \tau + \frac{\partial}{\partial t}\left( \nabla \cdot \boldsymbol{\Gamma} \right) = 4\pi G \varrho, \tag{9}$$

and

$$\nabla^2 \boldsymbol{\Gamma} - \frac{1}{c^2}\frac{\partial^2 \boldsymbol{\Gamma}}{\partial t^2} - \nabla\left( \nabla \cdot \boldsymbol{\Gamma} + \frac{1}{c^2}\frac{\partial \tau}{\partial t} \right) = \frac{4\pi G}{c^2}\boldsymbol{J}. \tag{10}$$

Equations (9) and (10) can be decoupled choosing the appropriate form of the potentials $\boldsymbol{\Gamma}$ and $\tau$. Moreover, if we simultaneously make the transformations

$$\boldsymbol{\Gamma} \to \boldsymbol{\Gamma}' = \boldsymbol{\Gamma} + \nabla \Lambda, \tag{11}$$

and

$$\tau \to \tau' = \tau - \frac{\partial \Lambda}{\partial t}. \tag{12}$$

in (5) and (8), we get the same original fields $\boldsymbol{g}$ and $\boldsymbol{K}$. Here, $\Lambda = \Lambda\left(x, y, z, t\right)$ is an arbitrary scalar function. We can choose this function in order to impose an additional condition over $\boldsymbol{\Gamma}$ and $\tau$, in a similar way like the Lorentz or Coulomb gauge in the electromagnetic field, namely,

$$\nabla \cdot \boldsymbol{\Gamma} = 0 \ \text{ or } \ \nabla \cdot \boldsymbol{\Gamma} = -\frac{1}{c^2}\frac{\partial \tau}{\partial t}, \tag{13}$$

which allows us to separate Equations (9) and (10) for the potentials $\tau$ and $\boldsymbol{\Gamma}$. These potentials depend on the gauge condition we chose.

## 3. Jefimenko's Equations for the Solenoidal and Irrotational Components

Following the ideas of the work of Chubykalo *et al.* [9]-[11] and using the analogy between the Maxwell's equations and the Jefimenko's ones [1] [2], we will apply the Helmholtz's theorem to define potentials that are independent of gauge transformations.

The Helmholtz's theorem claims that under certain conditions all vector fields can be represented as the sum of an irrotational and a solenoidal components. We will use this theorem to separate the fields $\boldsymbol{g}$ and $\boldsymbol{K}$.

Therefore, here we state the Helmholtz's theorem as [12]:

If the divergence $D(\boldsymbol{r})$ and a curl $\boldsymbol{C}(\boldsymbol{r})$ of a vector function $\boldsymbol{F}(\boldsymbol{r})$ are specified, and if they both go to zero faster than $1/r^2$ as $r \to \infty$, and if $\boldsymbol{F}(\boldsymbol{r})$ itself tends to zero as $r \to \infty$, then $\boldsymbol{F}(\boldsymbol{r})$ is uniquely given by

$$\boldsymbol{F} = -\nabla U + \nabla \times \boldsymbol{W}, \tag{14}$$

where

$$U(\boldsymbol{r}) = \frac{1}{4\pi} \iiint_{\text{All space}} \frac{D(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \mathrm{d}^3\boldsymbol{r}' \tag{15}$$

and

$$\boldsymbol{W}(\boldsymbol{r}) = \frac{1}{4\pi} \iiint_{\text{All space}} \frac{\boldsymbol{C}(\boldsymbol{r})}{|\boldsymbol{r} - \boldsymbol{r}'|} \mathrm{d}^3\boldsymbol{r}'. \tag{16}$$

We are going to suppose that all conditions of this theorem are satisfied by the fields $\boldsymbol{g}$ and $\boldsymbol{K}$ defined by Equations (1) to (4)[1], and then, we apply Helmholtz's theorem to these quantities, including $\boldsymbol{J}$. Thus, we obtain

$$\boldsymbol{g} = \boldsymbol{g}_i + \boldsymbol{g}_s, \tag{17}$$

$$\boldsymbol{K} = \boldsymbol{K}_i + \boldsymbol{K}_s, \tag{18}$$

$$\boldsymbol{J} = \boldsymbol{J}_i + \boldsymbol{J}_s, \tag{19}$$

where the indices "i" and "s" mean irrotational and solenoidal components of the vectors, respectively.

For example:

$$\boldsymbol{J}_i = -\frac{1}{4\pi} \nabla \iiint_{\text{All space}} \frac{\nabla' \cdot \boldsymbol{J}}{|\boldsymbol{r} - \boldsymbol{r}'|} \mathrm{d}^3\boldsymbol{r}', \tag{20}$$

$$\boldsymbol{J}_s = \frac{1}{4\pi} \nabla \times \iiint_{\text{All space}} \frac{\nabla' \times \boldsymbol{J}}{|\boldsymbol{r} - \boldsymbol{r}'|} \mathrm{d}^3\boldsymbol{r}', \tag{21}$$

We are going to substitute $\boldsymbol{g}, \boldsymbol{K}$ and $\boldsymbol{J}$ given by the Equations (17)-(19) into the Jefimenko's Equations (1)-(4) and then, we obtain for the irrotational part:

$$\nabla \cdot \boldsymbol{g}_i = -4\pi G \varrho, \tag{22}$$

$$\frac{\partial \boldsymbol{g}_i}{\partial t} = 4\pi G \boldsymbol{J}_i, \tag{23}$$

$$\nabla \cdot \boldsymbol{K}_i = 0, \tag{24}$$

$$\frac{\partial \boldsymbol{K}_i}{\partial t} = 0, \tag{25}$$

[1]For systems localized in a finite region of space, it is evident that the fields $\boldsymbol{g}$ y $\boldsymbol{K}$ depend on $r$ as $1/r^2$.

and the next equations for the solenoidal part:

$$\nabla \times \boldsymbol{g}_s = -\frac{\partial \boldsymbol{K}_s}{\partial t}, \tag{26}$$

$$\nabla \times \boldsymbol{K}_s - \frac{1}{c^2}\frac{\partial \boldsymbol{g}_s}{\partial t} = -\frac{4\pi G}{c^2}\boldsymbol{J}_s. \tag{27}$$

## 4. The Gravitodynamical Potentials from Helmholtz's Theorem

By definition, for the irrotational component of the gravitational field $\boldsymbol{g}_i$ we can define the scalar potential $\mathrm{T}$ as

$$\nabla \mathrm{T} = -\boldsymbol{g}_i \tag{28}$$

and if we substitute this relation into Equation (22), we obtain the Poisson's equation

$$\nabla^2 \mathrm{T} = 4\pi G \varrho. \tag{29}$$

Apparently, we need to take into account that $\mathrm{T}$ is not completely defined only by the Poisson's Equation (29), because we have another differential equation for $\mathrm{T}$, which can be obtained by substituting (28) into Equation (23)

$$\frac{\partial}{\partial t}\nabla \mathrm{T} = -4\pi G \boldsymbol{J}_i. \tag{30}$$

We show now that Equation (30) is equivalent to the law of conservation of mass. Indeed, let us take the divergence of the Equation (23), then we obtain as the result

$$\frac{\partial}{\partial t}\left(\nabla \cdot \boldsymbol{g}_i\right) = 4\pi G \nabla \cdot \boldsymbol{J}_i. \tag{31}$$

But from Equation (22) and because $\nabla \cdot \boldsymbol{J}_i = \nabla \cdot \left(\boldsymbol{J}_i + \boldsymbol{J}_s\right) = \nabla \cdot \boldsymbol{J}$, Equation (31) becomes the conservation mass law or the continuity equation

$$\nabla \cdot \boldsymbol{J} + \frac{\partial \varrho}{\partial t} = 0. \tag{32}$$

Now, we will demonstrate that the solution of Equation (30), indeed, is the same solution of the Poisson's Equation (29). To do this, we note that the irrotational component of $\boldsymbol{J}$ can be written as

$$\boldsymbol{J}_i = -\nabla \phi_J, \tag{33}$$

where the potential $\phi_J$ is defined as

$$\phi_J\left(x, y, z, t\right) = \frac{1}{4\pi}\iiint_{\text{All space}} \frac{\nabla' \cdot \boldsymbol{J}_i}{\left|\boldsymbol{r} - \boldsymbol{r}'\right|}\mathrm{d}^3\boldsymbol{r}' \tag{34}$$

or

$$\phi_J\left(x, y, z, t\right) = -\frac{1}{4\pi}\iiint_{\text{All space}} \frac{\frac{\partial}{\partial t}\varrho\left(\boldsymbol{r}', t\right)}{\left|\boldsymbol{r} - \boldsymbol{r}'\right|}\mathrm{d}^3\boldsymbol{r}', \tag{35}$$

and where, if we relate Equations (32), (23), (28) y (33) and the fact that $\nabla \cdot \boldsymbol{J}_i = \nabla \cdot \boldsymbol{J}$, we have

$$\frac{\partial(-\nabla \mathrm{T})}{\partial t} = 4\pi G(-\nabla \phi_J) \Rightarrow \frac{\partial \mathrm{T}}{\partial t} = 4\pi G \phi_J . \tag{36}$$

And from (36) and (34), we obtain

$$\frac{\partial \mathrm{T}}{\partial t} = -G \iiint_{\text{All space}} \frac{\frac{\partial}{\partial t} \varrho(\boldsymbol{r}', t)}{|\boldsymbol{r} - \boldsymbol{r}'|} \mathrm{d}^3 \boldsymbol{r}' \tag{37}$$

or

$$\mathrm{T} = -G \iiint_{\text{All space}} \frac{\varrho(\boldsymbol{r}', t)}{|\boldsymbol{r} - \boldsymbol{r}'|} \mathrm{d}^3 \boldsymbol{r}', \tag{38}$$

which is the solution of the Poisson's Equation (29). So we have found that the Poisson's equation given by Equation (29), completely defines the potential $\mathrm{T}$, together with its boundary conditions.

Since by definition $\boldsymbol{K} = \boldsymbol{K}_s$, then Equations (24) and (25) have the trivial solution $\boldsymbol{K}_i = 0$.

Let us now apply the Helmholtz's theorem to the vector potential $\boldsymbol{\Gamma}$. The Helmholtz theorem it is also known as the fundamental theorem of vector calculus (see Section III), and allows us to decompose every vectorial field in two components, an irrotational and a solenoidal one. Intuitively, it says that every vector function can be written as the sum of a divergence-free function (like $\boldsymbol{\Gamma}_s$) and a curl-free function (like $\boldsymbol{\Gamma}_i$), so that there exist scalar and vector potentials. So that $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_i + \boldsymbol{\Gamma}_s$, because this is the form in which we can easily solve the system formed by Equations (26) and (27). Supposing that

$$\boldsymbol{K}_s = \nabla \times \boldsymbol{\Gamma}_s \tag{39}$$

and taking into account (26) we obtain

$$\boldsymbol{g}_s = -\frac{\partial \boldsymbol{\Gamma}_s}{\partial t} \tag{40}$$

One can substitute Equations (39) and (40) into (27), and we obtain

$$\nabla^2 \boldsymbol{\Gamma}_s - \frac{1}{c^2} \frac{\partial^2 \boldsymbol{\Gamma}_s}{\partial t^2} = \frac{4\pi G}{c^2} \boldsymbol{J}_s, \tag{41}$$

where we used the vector identity $\nabla \times (\nabla \times \boldsymbol{V}) = \nabla(\nabla \cdot \boldsymbol{V}) - \nabla^2 \boldsymbol{V}$, for any arbitrary vector $\boldsymbol{V}$.

We have found that system of Equations (1)-(4) reduces to Equations (29) and (41), applying the Helmholtz's theorem. Therefore, we obtain separated equations for vector and scalar potentials, namely,

$$\nabla^2 \mathrm{T} = 4\pi G \varrho \tag{42}$$

and

$$\nabla^2 \boldsymbol{\Gamma}_s - \frac{1}{c^2} \frac{\partial^2 \boldsymbol{\Gamma}_s}{\partial t^2} = \frac{4\pi G}{c^2} \boldsymbol{J}_s. \tag{43}$$

## 5. Invariance of the Potentials $\Gamma_s$ and T under Gauge Transformations

Now, we will show that the potentials $\boldsymbol{\Gamma}_s$ and $\mathrm{T}$ (which are gravitodynamical countertypes of the electromagnetic potentials $\boldsymbol{A}_s$ and $\Phi$ from [9]) are invariant under gauge transformations (11) and (12), in common with the fields $\boldsymbol{g}$ and $\boldsymbol{K}$. This will be the most prominent property of these potentials.

If we apply the Helmholtz theorem to the gravitational and cogravitational fields in terms of the ordinary potentials given by (5) and (8) without taking into account any gauge condition, we have

$$\boldsymbol{g}_i = -\nabla\tau - \frac{\partial \boldsymbol{\Gamma}_i}{\partial t}, \tag{44}$$

$$\boldsymbol{g}_s = -\frac{\partial \boldsymbol{\Gamma}_s}{\partial t}, \tag{45}$$

$$\boldsymbol{K}_i = 0, \tag{46}$$

$$\boldsymbol{K}_s = \nabla \times \boldsymbol{\Gamma}_s, \tag{47}$$

then, by definition, $\boldsymbol{\Gamma}_i$ is given by

$$\boldsymbol{\Gamma}_i = -\nabla\phi_\Gamma, \tag{48}$$

where $\phi_\Gamma$ is an scalar function. If we substitute (48) into (44) then

$$\boldsymbol{g}_i = -\nabla\tau - \frac{\partial}{\partial t}\left(-\nabla\phi_\Gamma\right) = -\nabla\tau + \nabla\frac{\partial\phi_\Gamma}{\partial t} = -\nabla\left(\tau - \frac{\partial\phi_\Gamma}{\partial t}\right), \tag{49}$$

and from Equations (49) and (22) we have the relation between $\mathrm{T}$ and $\tau$.

$$\mathrm{T} = \tau - \frac{\partial\phi_\Gamma}{\partial t}. \tag{50}$$

Now, we apply the Helmholtz theorem and the gauge transformations (11) and (12) and from

$$\boldsymbol{\Gamma}' = \boldsymbol{\Gamma}'_i + \boldsymbol{\Gamma}'_s = \boldsymbol{\Gamma}_i + \boldsymbol{\Gamma}_s + \nabla\Lambda, \tag{51}$$

comparing the solenoidal parts we obtain

$$\boldsymbol{\Gamma}'_s = \boldsymbol{\Gamma}_s. \tag{52}$$

If we seek the transformation law for $\phi_\Gamma$, then we can obtain the other transformation law for $\mathrm{T}$.

From Equation (51), we have the irrotational part of $\boldsymbol{\Gamma}'$

$$\boldsymbol{\Gamma}'_i = \boldsymbol{\Gamma}_i + \nabla\Lambda, \tag{53}$$

and including Equation (48) in (53) we get

$$-\nabla\phi'_\Gamma = -\nabla\phi_\Gamma + \nabla\Lambda = -\nabla\left(\phi_\Gamma - \Lambda\right) \tag{54}$$

or

$$\phi'_\Gamma = \phi_\Gamma - \Lambda. \tag{55}$$

At last, we can use Equation (50) for $\mathrm{T}$ and, if we consider Equations (11), (12) and

(55), we obtain

$$\mathrm{T}' = \tau' - \frac{\partial \phi'_{\Gamma}}{\partial t} = \left(\tau - \frac{\partial \Lambda}{\partial t}\right) - \frac{\partial}{\partial t}\left(\phi_{\Gamma} - \Lambda\right) = \tau - \frac{\partial \phi_{\Gamma}}{\partial t} = \mathrm{T}. \tag{56}$$

Hence, we have checked that $\mathbf{\Gamma}_s$ and $\mathrm{T}$ are invariants under gauge transformation and we can see that any gauge transformation is irrelevant if we use the Helmholtz's theorem.

It is convenient to remark that the fields $\boldsymbol{g}$ and $\boldsymbol{K}$ are generated only by $\mathbf{\Gamma}_s$ and $\mathrm{T}$ given by (28), (39) and (40), so we can consider $\mathbf{\Gamma}_s$ and $\mathrm{T}$ as the potentials generating the gravitodynamical field $\boldsymbol{g}$ and $\boldsymbol{K}$.

## 6. Conclusions

And so, we have shown that it is possible to define vector as well as scalar gravitodynamical potentials, which are invariant under gauge transformation. These potentials are defined uniquely from their differential Equations (42) and (43). For this reason, we have arguments for supposing the physical reality of these potentials, similarly to the fields $\boldsymbol{g}$ and $\boldsymbol{K}$ and unlike the gravitational potentials introduced by Jefimenko in [1], which are only used as a mathematical tool for calculating gravitational and cogravitational fields.

Our scalar potential $\mathrm{T}$ is a generator of the so-called instantaneous action at a distance in gravitation, and the vector potential $\mathbf{\Gamma}_s$ can propagate with the velocity of light and it is responsible for the retarded action of the gravitodynamical field. So, one can conclude that retarded interaction in gravitodynamics takes place *not instead* but *together with* instantaneous action at a distance.

## References


[1] Jefimenko, O.D. (2006) Gravitation and Cogravitation. Electret Scientific Company, Star City.

[2] Jefimenko, O.D. (2000) Causality Electromagnetic Induction and Gravitation: A Different Approach to the Theory of Electromagnetic and Gravitational Fields. 2nd Edition, Electret Scientific Company, Star City.

[3] Heaviside, O. (1893) *The Electrician*, **31** 281-282.

[4] Brillouin, L. (1970) Relativity Reexamined. Academic Press, New York.

[5] Einstein, A. (1912) *Vierteljahrschrift für gerichtliche Medizin und öffentliches Sanitatswessen*, **44**, 1-4.

[6] González, W.S. (2013) Gravitoelectromagnetismo y principio de Mach, eWT Ediciones.

[7] Vetö, B. (2010) *European Journal of Physics*, **31**, 5.

[8] Vetö, B. (2011) *European Journal of Physics*, **32**, 5.

[9] Chubykalo, A., Espinoza, A. and Alvarado Flores, R. (2011) *Physica Scripta*, **84**, 015009-1-6. http://dx.doi.org/10.1088/0031-8949/84/01/015009

[10] Chubykalo, A., Espinoza, A. and Alvarado Flores, R. (2012) *Physica Scripta*, **85**, 047002. http://dx.doi.org/10.1088/0031-8949/84/6/069502

[11] Chubykalo, A., Espinoza, A., Alvarado Flores, R. and Gutiérrez Rodríguez, A. (2006)

*Foundations of Physics Letteres*, **19**, 37-49.
http://dx.doi.org/10.1088/0031-8949/85/04/047002

[12] Arfken, G.B. and Weber, H.J. (1995) Mathematical Methods for Physicists. Academic Press, New York.

Scientific Research Publishing

Journal of Modern Physics, 2016, 7, 1627-1634
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Gauge Invariance, the Quantum Metric Tensor and the Quantum Fidelity

**J. Alvarez-Jiménez, Jose David Vergara**

Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Ciudad de México, México
Email: vergara@nucleares.unam.mx

**How to cite this paper:** Alvarez-Jiménez, J. and Vergara, J.D. (2016) Gauge Invariance, the Quantum Metric Tensor and the Quantum Fidelity. *Journal of Modern Physics*, **7**, 1627-1634.
http://dx.doi.org/10.4236/jmp.2016.713147

**Received:** June 3, 2016
**Accepted:** September 3, 2016
**Published:** September 6, 2016



## Abstract

The quantum metric tensor was introduced for defining the distance in the parameter space of a system. However, it is also useful for other purposes, like predicting quantum phase transitions. Due to the physical information this tensor provides, its gauge independence sounds reasonable. Moreover, its original construction was made by looking for this gauge independence. The aim of this paper, however, is to prove that the quantum metric tensor does depend on the gauge. In addition, a real gauge invariant quantum metric tensor is introduced. A related concept is the quantum fidelity, which is also shown to depend on the gauge in this paper. The gauge dependences are explicitly shown by computing the quantum metric tensor and the quantum fidelity of the Landau problem in different gauges. Then, a real gauge independent metric tensor is proposed and computed for the same Landau problem. Since the gauge dependences have not been observed before, the results of this paper might lead to a new study of topics that are believed to be completely understood.

## Keywords

Landau problem, Quantum Metric Tensor, Gauge Dependence, Quantum Fidelity

## 1. Introduction

The main purpose for constructing the quantum metric tensor (QMT) was to define a distance in the system's parameter space [1] and recently it has been shown that this metric tensor can be obtained using the renormalization flow equations [2]. That is why it is not surprising that the QMT is related to the quantum fidelity (QF), which is also used for measuring the distance between states [3], even though some studies have shown that the QMT can also be used to predict quantum phase transitions [4] [5]. In [6], the critical exponents for systems that present continuous second order phase transitions are defined. Moreover, the geodesics induced by the QMT have been useful for analyzing the phase transitions [7]. In general, the Riemannian structure introduced by the QMT has been studied in some particular systems, seen for example [7]. The authors of [8] make an analysis of the Gaussian curvature induced by the QMT and describe the critical phenomena in relation with this curvature.

In general, there has been much interest in the geometrical properties of quantum systems. In [9] it is shown that the mass can be seen as a geometric effect in the Hilbert space. Reference [10] proposes a formalism of quantum space geometry for generalized coherent states and analyzes it with known results of the symmetry AdS/CFT. On the other hand, in Ref. [11] a numerical analysis of the fractional quantum Hall effect related with geometric stability is performed. Continuing with the numerical computation, in [12], it is presented a method to compute the fidelity susceptibility (a particular case of the QMT) with the Monte Carlo method.

The QMT was constructed by looking for a gauge independence [1] and, in fact, it was partially done. However, when we consider some kinds of gauge transformations, the QMT is not invariant. Nevertheless, in current works the gauge dependence is overtly assumed. Since this gauge dependence has not been observed before, we explain its origin and propose a real gauge invariant quantum metric tensor. On the other hand, the QF is a similar concept that is also useful to measure a distance in the parameter's space of a system, and as well as the QMT, it sounds reasonable that it does not depend on the gauge. However, it is proved that it is not always the case.

In this paper, we use the Landau problem to show the gauge dependence of the QMT and the QF. For this reason, in Section 2 we describe the Landau problem in the symmetric gauge. Section 3 shows the QMT for one of the ground states in different gauges. While Section 4 introduces a gauge independent definition of the QMT, Section 5 shows the calculation of this new definition for the Landau problem. On the other hand, Section 6 shows the gauge dependence of the QF and explains its origin. Finally, a discussion and our conclusions are written in Section 6.

## 2. The Landau Problem

The Landau problem [13] consists on a charged particle interacting with a constant and homogeneous magnetic field, $\boldsymbol{B}$. If we consider a particle of unitary mass and charge $e<0$ the Hamiltonian of the system is given by

$$H=\frac{1}{2}\left(\boldsymbol{P}-\frac{e}{c}\boldsymbol{A}\right)^2, \tag{1}$$

where $\boldsymbol{A}$ is the vector potential, such that $\boldsymbol{B}=\nabla\times\boldsymbol{A}$. If we assume that the magnetic field points in the $z$ direction *i.e.* $\boldsymbol{B}=B\hat{z}$, then the movement in $z$ will be constant, and we can ignore it. For the quantum case, the energy spectrum of the Landau problem is given by [13]

$$E_n=\hbar\omega\left(n+\frac{1}{2}\right),\quad n=0,1,2,\cdots,\quad \omega=\frac{|eB|}{c}, \tag{2}$$

these $E_n$ are the well-known Landau levels. However, for the Landau problem each level is infinitely degenerated. Therefore, we need an additional Hermitian operator, which commutes with $H$, to label the states. If we choose the symmetric gauge, *i.e.*

$$\boldsymbol{A}_s = \frac{B}{2}(-y, x), \tag{3}$$

we can select the angular momentum in the $z$ direction, $L_z = xp_y - yp_x$, as the second operator. In this gauge, the ground states are given by [14]

$$\psi_{o,m} = \sqrt{\frac{1}{\pi m!}\left(\frac{eB}{2\hbar c}\right)^{m+1}} \left(x + iy\right)^m \mathrm{e}^{-\frac{eB}{4\hbar c}\left(x^2+y^2\right)}, \tag{4}$$

where $m$ is a label for the angular momentum in the $z$ direction, such that

$$L_z \psi_{0,m} = \hbar m \psi_{0,m} \tag{5}$$

In this case, we see that the wavefunction depends on the parameters space and the physical space $x$.

## 3. The Quantum Metric Tensor of the Landau Problem

The QMT, $G_{ij}$, is useful to define a distance in the system's parameter space [1]. If our quantum system depends on $n$ parameters, $\lambda_i$, the QMT is given by

$$G_{ij} = \Re e\left[\left(\partial_i \psi, \partial_j \psi\right)\right] - \beta_i \beta_j, \tag{6}$$

where $\psi$ is the state of the system, $\partial_i = \frac{\partial}{\partial \lambda_i}$ and

$$\beta_i = -i\left(\psi, \partial_i \psi\right), \tag{7}$$

with this definition, the corresponding distance will be [1]

$$\mathrm{d}l^2 = G_{ij}\mathrm{d}\lambda_i \mathrm{d}\lambda_j. \tag{8}$$

It is proved that the QMT is gauge invariant [1], nevertheless, this proof is not the most general. The demonstration assumes some specific features of the phase difference caused by a gauge transformation. In order to show the gauge dependence of the QMT, we need to compute it in different gauges. The first calculations will be in the symmetric gauge.

### 3.1. The Quantum Metric Tensor in the Symmetric Gauge

For the purpose of this paper, it is sufficient to consider only the variation of $B$, therefore the parameter space will be 1-dimensional, with $\lambda_1 = B$, and setting $G_{BB} = G$ is appropriate. We will compute the QMT of the ground state with $m = 0$, then, by using the state presented in Equation (4), the first term of the definition will be

$$\Re e\left[\left(\partial_B \psi, \partial_B \psi\right)\right] = \frac{1}{2B}, \tag{9}$$

whereas

$$\beta_B = 0, \tag{10}$$

therefore

$$G = \frac{1}{2B}. \tag{11}$$

### 3.2. Comparison of the Quantum Metric Tensor in Different Gauges

In order to prove the gauge dependence of the QMT, we make the calculation in different gauges. It is known [15] that when two gauges are related by

$$\boldsymbol{A}_2 = \boldsymbol{A}_1 + \nabla\Lambda(\lambda, \boldsymbol{x}), \tag{12}$$

the corresponding wave functions obey

$$\psi_2(\lambda, \boldsymbol{x}) = \exp\left(i\frac{e}{\hbar c}\Lambda\right)\psi_1(\lambda, \boldsymbol{x}). \tag{13}$$

According to the theory [1] [16], since the wave functions are related just by a change of phase, the QMTs should coincide. To explicitly show that this match does not always occur, we choose $\boldsymbol{A}_1$ as the symmetric gauge and

$$\Lambda = gBxy. \tag{14}$$

This particular $\Lambda$ allows us to examine several gauges using $g$ as a parameter. In particular, when we set $g = 1/2$, we obtain the Landau Gauge $\boldsymbol{A}_L$ given by

$$\boldsymbol{A}_L = B(0, x), \tag{15}$$

and with $g = 0$ we recover the symmetric gauge. In this case, $\Lambda$ depends on the parameter $B$ and the physical space $(x, y)$.

Now, in Equation (13), we set $\psi_1$ as the ground state in the symmetric gauge, then the ground state with $m = 0$ in the new gauge will be

$$\psi'_{0,0}(B, x, y) = \sqrt{\frac{eB}{2\pi\hbar c}}\exp\left(-\frac{eB}{4\hbar c}\left(x^2 + y^2\right)\right)\exp\left[i\frac{eg}{\hbar c}Bxy\right]. \tag{16}$$

From Equation (16) and the definition of the QMT, we compute that

$$G' = \left(g^2 + \frac{1}{2}\right)\frac{1}{B}. \tag{17}$$

The presence of $g$ in Equation (17) clearly implies gauge dependence. This gauge dependence is inherited by the distance in the parameter space. This means that we do not have a gauge independent distance in the parameter space. In the specific case of Equation (17) the distance is minimum when we work in the symmetric gauge ($g = 0$), and it increases indefinitely as we increase the absolute value of $g$. It is worth to notice, however, that the QMT diverges when $B \to 0$ for any value of $g$.

## 4. Real Gauge Invariant Quantum Metric Tensor

If we perform a gauge transformation in the parameter space, given by

$$\psi' = \mathrm{e}^{i\alpha(\lambda, x)}\psi, \tag{18}$$

then $\beta_i$ changes as

$$\beta'_i = \beta_i + \left(\psi, (\partial_i\alpha)\psi\right). \tag{19}$$

It has been assumed that the phase $\alpha$, as well as its derivatives, can be taken outside of the internal product. Therefore, we would be able to simplify Equation (19) to

$$\beta_i' = \beta_i + \partial_i \alpha, \tag{20}$$

when Equation (20) is valid, the tensor presented in Equation (6) is gauge invariant. This means that the QMT is gauge invariant when $\partial_i \alpha$ is independent of the measure of the internal product.

However, some phases, and its derivatives, may depend on the physical space, $(x, y)$, or any other operators. See, for example, the phase in Equation (16). In these cases, Equation (19) cannot be simplified; thus, the tensor of Equation (6) is no longer gauge invariant. It is worth to notice that Equation (12) and Equation (20) seem to give the same transformation rule. However, in Equation (12) the derivatives are computed respect to the coordinates, while in Equation (20), one derives respect to the parameters $\lambda_i$.

Before constructing the real gauge invariant QMT, we note that Equation (6) can be written as

$$G_{ij} = \Re e\left[\left(\left(\partial_i - i\beta_i\right)\psi, \left(\partial_j - i\beta_j\right)\psi\right)\right], \tag{21}$$

or, in the representation of coordinates

$$G_{ij} = \Re e\left[\int \mathrm{d}^3 x \left(\partial_i + i\beta_i\right)\psi^* \left(\partial_j - i\beta_j\right)\psi\right], \tag{22}$$

because $\beta_i$ is real. Then, by looking Equation (20), we realize that $\beta_i$ transforms like a connection when $\alpha$ is independent of the internal product. This means that the QMT of Equation (21) is constructed with covariant derivatives, using $\beta_i$ as the connection. Nonetheless, in the general case $\beta_i$ transforms like it is shown in Equation (19), and it cannot be used as the connection.

For constructing the gauge invariant QMT, we need a function $\Gamma_i$ that transforms like

$$\Gamma_i' = \Gamma_i + \partial_i \alpha, \tag{23}$$

when we perform a change of gauge given by Equation (18). With this new connection, the gauge invariant QMT will be

$$G_{ij} = \Re e\left[\left(\left(\partial_i - i\Gamma_i\right)\psi, \left(\partial_j - i\Gamma_j\right)\psi\right)\right], \tag{24}$$

or

$$G_{ij} = \Re e\left[\int \mathrm{d}^3 x \left(\partial_i + i\Gamma\right)\psi^* \left(\partial_j - i\Gamma_j\right)\psi\right]. \tag{25}$$

In Equations (24) and (25), we recognize the covariant derivative, $D_i$, given by

$$D_i = \partial_i - i\Gamma_i, \tag{26}$$

which transforms like

$$\left(D_i \psi\right)' = \mathrm{e}^{i\alpha} D_i \psi, \tag{27}$$

under a change of gauge. Using the covariant derivative, the QMT takes the form

$$G_{ij} = \Re e\left(D_i\psi, D_j\psi\right) = \Re e\left[\int \mathrm{d}^3x\left(D_i\psi\right)^* D_j\psi\right]. \tag{28}$$

Equation (28) defines a gauge invariant QMT. Since Equation (27) is valid, then Equation (28) will always be gauge independent. Here we can see that if, instead of Equation (27), we perform a no Abelian gauge transformation we would generalize the QMT to a no abelian QMT.

However, we need to find the correct connection that transforms like it is shown in Equation (23). The form of the new connection $\Gamma_i$ will depend on the specific problem to be analyzed. $\Gamma_i$ must reduce to $\beta_i$ in the case that $\partial_i\alpha$ can be taken outside of the internal product. Therefore, the new QMT must also reduce to the one presented in Equation (6) when $\partial_i\alpha$ is independent of the measure of the internal product. In the following section we present the $\Gamma_B$ for the example studied in this paper.

## 5. Gauge Invariant Quantum Metric Tensor of the Landau Problem

Continuing with the example presented in Section 3.2, the new QMT is given by

$$G = \Re e\left(\left(\partial_B - i\Gamma_B\right)\psi, \left(\partial_B - i\Gamma_B\psi\right)\right). \tag{29}$$

The fact that $\beta_B = 0$ in the usual case, suggests that we must set

$$\Gamma_B = 0, \tag{30}$$

therefore, according to Equation (23), and using that $\alpha = \frac{eg}{\hbar c}Bxy$, we find

$$\Gamma'_B = \frac{gexy}{\hbar c}, \tag{31}$$

thus, under the transformation of Equation (18), we get

$$G' = \Re e\left[\int \mathrm{d}^3x\left(\partial_B + i\frac{gexy}{\hbar c}\right)\left(\psi'\right)^*\left(\partial_B - i\frac{gexy}{\hbar c}\right)\psi'\right]. \tag{32}$$

Applying Equation (32) to the state given by Equation (16), we obtain

$$G' = \frac{1}{2B}, \tag{33}$$

for any gauge. That is, the QMT proposed in this paper is gauge independent.

## 6. Gauge Dependence of the Quantum Fidelity

As it was mentioned in the introduction, the QF is also useful to measure the distance between states. If the quantum system depends on $n$ parameters $\lambda$, the QF is given by [2]

$$F\left(\lambda, \lambda'\right) = \left|\left(\psi\left(\lambda\right), \psi\left(\lambda'\right)\right)\right|, \tag{34}$$

where $\psi$ is the state of the system. Continuing with the Landau problem, if we compute the QF using the state given by (16), we obtain that

$$F\left(B, B'\right) = \frac{2\sqrt{BB'}}{\sqrt{\left(B+B'\right)^2 + 4g^2\left(B-B'\right)}}. \tag{35}$$

In Equation (35) we can see dependence with the parameter *g*, therefore the QF also depends on the gauge chosen. This dependence occurs for the same reason that it appears in the QMT *i.e*. the phase difference is not independent of the internal product. If we start with a gauge whose state vector is $\psi_1$ the QF will be

$$F(\lambda,\lambda') = \left|\left(\psi_1(\lambda),\psi_1(\lambda')\right)\right|, \tag{36}$$

if we now perform the gauge transformation given by Equation (18), the QF fidelity will take the form

$$F(\lambda,\lambda') = \left|\left(\mathrm{e}^{i\frac{e}{\hbar c}\Lambda(\lambda)}\psi_2(\lambda),\mathrm{e}^{i\frac{e}{\hbar c}\Lambda(\lambda')}\psi_2(\lambda')\right)\right|, \tag{37}$$

again, when the phase can be taken outside of the internal product Equation (37) simplifies to

$$F(\lambda,\lambda') = \left|\left(\psi_2(\lambda),\psi_2(\lambda')\right)\right| \tag{38}$$

and the fidelities in different gauges coincide. However, for more general gauges, like the presented in the example studied here, we cannot take the phase outside and, therefore, the fidelities do not coincide.

## 7. Discussion and Conclusions

We explicitly showed that the QMT and the QF depend on the gauge. This dependence is directly related to the phase difference between the wave functions in different gauges: when the change of gauge introduces a phase whose derivatives $\partial_i\alpha$ can be taken outside of the internal product, both, the QMT and the QF are invariant. However, when general phases are considered, they depend on the gauge.

We also proposed a real gauge invariant QMT by defining a new connection $\Gamma_i$ that transforms according to Equation (23). Despite the gauge independence, the connection $\Gamma_i$ was not explicitly given, and the form of $\Gamma_i$ will depend on the specific problem to be studied. In the example shown in this paper, *i.e*. the Landau Problem, we successfully proposed the correct $\Gamma_i$ for obtaining a gauge invariant QMT.As it was pointed out before, the QMT and the QF have several applications in physics. The fact that the QMT and the QF depend on the gauge can give rise to new studies in the topics that apply this two tools. An important case is the applicability of the QMT for predicting quantum phase transitions. In the example studied in this paper, the gauge independent QMT, as well as the gauge dependent QMT, diverges for the same value of the field *B* in any gauge. This fact suggests that both QMTs are useful for predicting quantum phase transitions. However, the chosen gauge in this example is not the most general, and further studies are necessary.

## Acknowledgements

This work was partially supported by DGAPA-PAPIIT grant IN103716; CONACyT project 237503, and scholarship 419420. We also wish to acknowledge Unidad de Posgrado, UNAM for the support and the workshop "Academic Writing" during the

preparation of this paper.

## References


[1] Provost, J.P. and Vallee, G. (1980) *Communications on Mathematical Physics*, **76**, 289-301. http://dx.doi.org/10.1007/BF02193559

[2] Maity, R., Mahapatra, S. and Sarkar T. (2015) *Physical Review E*, **92,** 052101-1-052101-10. http://dx.doi.org/10.1103/PhysRevE.92.052101

[3] Gu, S.-J. (2010) *International Journal of Modern Physics B*, **24**, 4371-4458. http://dx.doi.org/10.1142/S0217979210056335

[4] Zanardi, P., Giorda, P. and Cozzini, M. (2007) *Physical Review Letters*, **99**, 100603-1-100603-4. http://dx.doi.org/10.1103/PhysRevLett.99.100603

[5] Campos Venuti, L. and Zanardi, P. (2007) *Physical Review Letters*, **99**, 095701-1-095701-4. http://dx.doi.org/10.1103/PhysRevLett.99.095701

[6] Kumar, P. and Sarkar, T. (2014) *Physical Review E,* **90**, 042145-1-042145-8. http://dx.doi.org/10.1103/PhysRevE.90.042145

[7] Kumar, P., Mahapatra, S., Phukon, P. and Sarkar, T. (2012) *Physical Review E,* **86**, 051117-1-051117-6. http://dx.doi.org/10.1103/PhysRevE.86.051117

[8] Kolodrubetz, M., Gritsev, V. and Polkovnikov, A. (2013) *Physical Review B*, **88**, 064304-1-064304-17. http://dx.doi.org/10.1103/PhysRevB.88.064304

[9] D'Alessio, L. and Polkovnikov, A. (2014) *Annals of Physics,* **345**, 141-165. http://dx.doi.org/10.1016/j.aop.2014.03.009

[10] Kriel, J.N., van Zyl, H.J.R. and Scholtz, F.G. (2015) *Journal of High Energy Physics,* **2015**, 1-26. http://dx.doi.org/10.1007/JHEP11(2015)140

[11] Jackson, T.S., Möller, G. and Roy, R. (2015) *Nature Communications,* **6**, 1-11. http://dx.doi.org/10.1038/ncomms9629

[12] Wang, L., Liu, Y.-H., Imriška, J., Ma, P.N. and Troyer, M. (2015) *Physical Review X,* **5**, 031007-1-031007-14. http://dx.doi.org/10.1103/PhysRevX.5.031007

[13] Landau, L.D. and Lifshitz, E.M. (1977) Quantum Mechanics Non-Relativistic Theory. Pergamon Press, Oxford.

[14] Haugset, T., Ruud, J.A. and Ravndal, F. (1993) *Physica Scripta,* **47**, 715-719. http://stacks.iop.org/1402-4896/47/i=6/a=004

[15] Sakurai, J.J. and Napolitano, J. (2011) Modern Quantum Mechanics. Addison-Wesley, San Francisco.

[16] Chruscinski, D. and Jamiolkowski, A. (2004) Geometric Phases in Classical and Quantum Mechanics. Springer Science + Business Media, New York.

Journal of Modern Physics, 2016, 7, 1635-1643
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Electron-Domain Wall Interaction with a Ferromagnetic Spherical Domain Wall

**Leonardo dos Santos Lima**

Departamento de Física e Matemática, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Brazil
Email: lslima@des.cefetmg.br

**How to cite this paper:** Lima, L.S. (2016) Electron-Domain Wall Interaction with a Ferromagnetic Spherical Domain Wall. *Journal of Modern Physics*, **7**, 1635-1643.
http://dx.doi.org/10.4236/jmp.2016.713148

**Received:** August 10, 2016
**Accepted:** September 3, 2016
**Published:** September 6, 2016



## Abstract

The interaction between an electron with a three-dimensional domain wall was investigated using the Born's expansion of the $S$ scattering matrix. We obtain an influence of the scattering of the electron with the ferromagnetic domain wall in the spin wave function of the electron with the aim to generate the knowledge about the state of the electron spin after the scattering. It relates to the recent problem of generation of the spin polarized electric current. We also obtain the contribution of the electron-wall domain interaction on the electric conductivity $\sigma(\omega)$, through the wall domain, where we have obtained a peak of resonance in the conductivity for one value of $\omega \simeq 2.5J$.

## Keywords

Electron Scattering, Domain-Wall, Ferromagnet

## 1. Introduction

The study of the quantum matter is an important topic of research in modern physics [1]. The description of particles interacting at low temperature is crucial in the determining and distinguishing of characteristics being an important problem in condensed matter physics. One of these problems is the study of the $s-d$-electron-scattering with the ferromagnetic domain wall [2], where an electric current density crosses a ferromagnetic metallic film. The domain wall resistivity is a rather old topic and has been thoroughly studied by many research groups [3]-[5].

The injection of a spin current in a magnetic film can generate a spin-transfer torque that acts on the magnetization collinearly to the damping torque [6]-[9]. Recently, the electron scattering by an one-dimensional domain wall in quantum wires that were described by the Luttinger liquid model was studied using Bosonization and Renormalization group [10] [11]. The transport and the scattering in quantum wires with domain wall were considered in [12]. The interaction of the domain wall with an interacting one-dimensional electron gas was studied in [13]. Peter *et al.* [14] have studied the influence of the domain wall scattering in the electron resistivity. Moreover, the importance to study the influence of scattering electron-domain wall is due the connection with phenomena depending on the spin such as the Giant Magneto-Resistance (GMR) [15], where we can have a large variation of the electric resistance with the variation of magnetization through the domain wall. The Quantum information technology promises one faster and more secure means of data manipulation that makes use of the quantum properties of the matter [16]-[18]. This demands the control of the spin of the electron and the needless of filtration of the electric current. We transform an electric current spin polarized by the interaction of the spins of the electrons that compose the electric current with the spins of the wall domain, $V(\boldsymbol{r})$ [19]-[21].

The model that we are interesting is described by the following Hamiltonian

$$\mathcal{H} = -\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r}) + V(\boldsymbol{r})\psi(\boldsymbol{r}) + J\hat{s}\cdot\hat{S}(\boldsymbol{r})\psi(\boldsymbol{r}). \tag{1}$$

where $J$ denotes the exchange interaction, $V(\boldsymbol{r})$ is the nonmagnetic periodic potential of the lattice and the last term $J\boldsymbol{s}\cdot\boldsymbol{S}(\boldsymbol{r}')$ represents the potential of scattering of a electron spin with the spins of the domain wall. In an homogeneous magnetic domain wall, the magnetization is collinear, $\hat{S}(r) = \hat{S}$, hence it is natural to choose this direction for the axis of quantization of the spin of electron $\boldsymbol{s}$. The interaction of each electron with the spins of the wall is depicted in **Figure 1**. In **Figure 2**, we present the behavior of the potential of interaction between the spin of the electron with the spins of the wall domain.

The purpose of this paper is to verify the influence of the scattering of electrons with the ferromagnetic domain wall on the spin wave function of the electron. We have employed the Borns approximation and the Matsubara's Green's function method to study the influence of the scattering of electron with the wall domain. The paper is divided in the following way. In Section 2, we discuss about the mechanism of electron scattering with the domain wall, in the Section 3, we verify the influence of electron-wall domain interaction in the current and finally, in the last section, Section 4, we present the final remarks.

## 2. Electron Scattering with the Domain Wall

We use the Born's expansion of $f(k,k')$ to calculate the influence of the domain wall on the spin wave function of the electron. In large distance of the wall domain, the state of the electron is given by

$$\psi(\boldsymbol{r}) = \begin{pmatrix} \phi_\uparrow(\boldsymbol{r}) \\ \phi_\downarrow(\boldsymbol{r}) \end{pmatrix}, \tag{2}$$

where $\psi_{in}(\boldsymbol{r}) = \psi(-\infty)$ is the state of the electron after interacting with the wall. In large distance of the domain wall, $\boldsymbol{r}$, the eigenstates $\psi_{out}(\boldsymbol{r})$ are given by

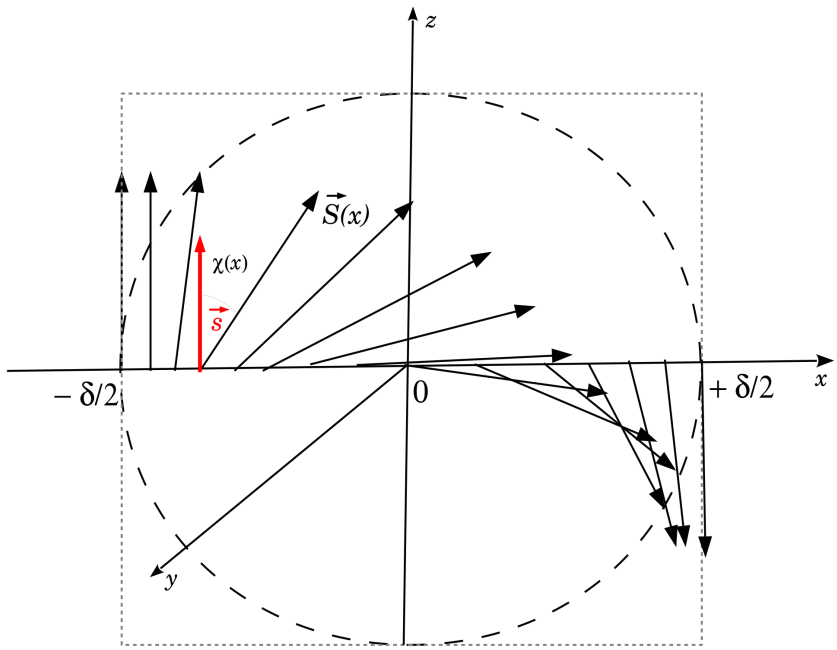



**Figure 1.** Interaction between one electron with the spins of the spherical domain wall of radius $\delta/2$.

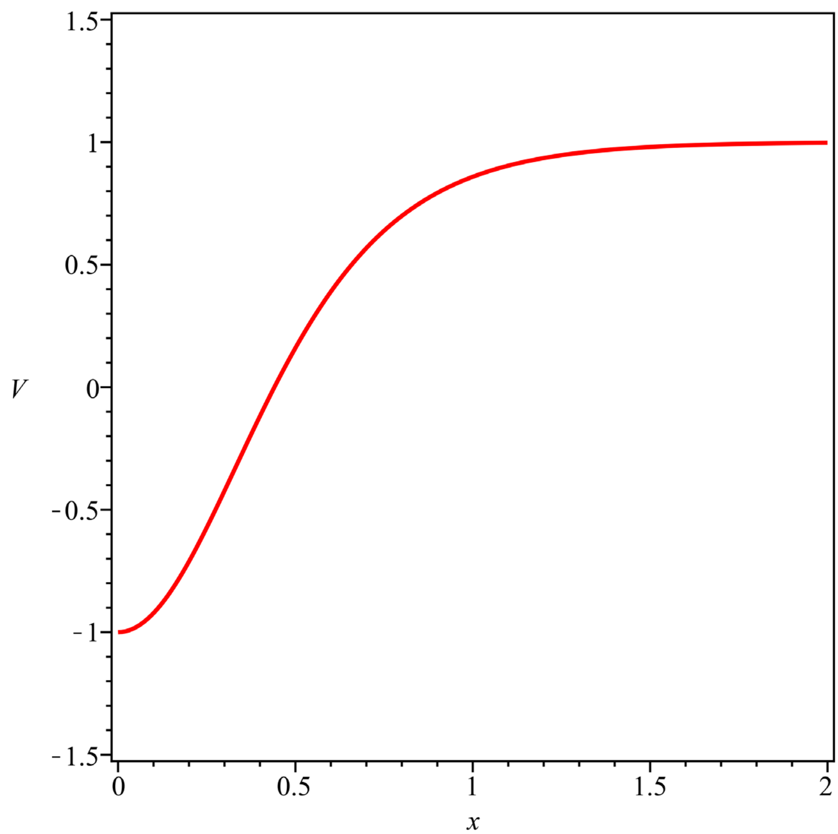



**Figure 2.** Behavior of the potential of interaction between the electron with the domain wall, $V(x)$, where the width of the wall is $\delta = 2$.

$$\psi_{\uparrow}(\boldsymbol{k},\boldsymbol{r}) = \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}\begin{pmatrix}1\\0\end{pmatrix} + \frac{\mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{r} f(k,k')\begin{pmatrix}0\\1\end{pmatrix}$$
$$\psi_{\downarrow}(\boldsymbol{k},\boldsymbol{r}) = \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}\begin{pmatrix}0\\1\end{pmatrix} + \frac{\mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{r} f(k,k')\begin{pmatrix}1\\0\end{pmatrix} \quad (3)$$

where $|\psi(\boldsymbol{r})\rangle = \boldsymbol{S}|\psi(\boldsymbol{r})\rangle$ and $\boldsymbol{S}$ is the scattering matrix.

$$f(k,k') = -\frac{\mu}{2\pi\hbar^2}\langle \boldsymbol{k}|\boldsymbol{T}|\boldsymbol{k}'\rangle, \quad (4)$$

as $\boldsymbol{T}$ is the transition matrix, given by the Lipmann-Schwinger's equation

$$\boldsymbol{T} = \boldsymbol{V} + \boldsymbol{V}\frac{1}{\omega - \mathcal{H}_0 + i0^+} \quad (5)$$

and

$$\mathcal{H}_0 = -\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}), \quad (6)$$

$$V(\boldsymbol{r}) = J\hat{s}\cdot\hat{S}(\boldsymbol{r}). \quad (7)$$

We consider $J = 1$.

$$f(k,k') = \sum_{n=0}^{\infty} f^{(n)}(k,k'); \quad (8)$$

as $n$ being the number of times that the $\boldsymbol{V}$ operators enters,

$$f^{(1)}(k,k') = -\frac{\mu}{2\pi^2\hbar}\langle \boldsymbol{k}|\boldsymbol{V}|\boldsymbol{k}'\rangle \quad (9)$$

and

$$V(x') = -\frac{2J_{sd}}{g\mu_B}\boldsymbol{s}\cdot\langle \hat{S}(x')\rangle, \quad (10)$$

where $x'$ corresponds to the region of $x$ into the ferromagnetic wall domain. We have that

$$f^{(1)}(k,k') = -\frac{2}{e}A(k,k') \quad (11)$$

where $e$ is the electron charge and $A(k)$ is given by

$$A(k) = \int_{\frac{-\delta}{2}}^{+\frac{\delta}{2}} S\cos\left(\arctan\left(\mathrm{e}^{-\delta x'}\right)\right)\sin(kx')\,\mathrm{d}x'. \quad (12)$$

The integral in the Equation (12) was solved approximately using the Maple program as

$$A(k)\simeq-\frac{1}{4\left(k^2+\delta^2\right)}\Bigg[S\sqrt{2}\Bigg(\mathrm{e}^{\frac{\delta^2}{2}}k\cos\left(\frac{\delta k}{2}\right)\delta^3+\mathrm{e}^{\frac{\delta^2}{2}}k^3\cos\left(\frac{\delta k}{2}\right)\delta-4\mathrm{e}^{\frac{\delta^2}{2}}k\cos\left(\frac{\delta k}{2}\right)\delta$$
$$+\mathrm{e}^{\frac{\delta^2}{2}}\sin\left(\frac{\delta k}{2}\right)\delta^4-\mathrm{e}^{\frac{\delta^2}{2}}k\sin\left(\frac{\delta k}{2}\right)\delta^2k^2+2\mathrm{e}^{\frac{\delta^2}{2}}\sin\left(\frac{\delta k}{2}\right)\delta^2-2\mathrm{e}^{\frac{\delta^2}{2}}k\sin\left(\frac{\delta k}{2}\right)k^2 \tag{13}$$
$$+\mathrm{e}^{\frac{-\delta^2}{2}}k\cos\left(\frac{\delta k}{2}\right)\delta^4+\mathrm{e}^{-\frac{\delta^2}{2}}k^3\cos\left(\frac{\delta k}{2}\right)\delta+\mathrm{e}^{\frac{-\delta^2}{2}}k\cos\left(\frac{\delta k}{2}\right)\delta^4+\mathrm{e}^{-\frac{\delta^2}{2}}k\sin\left(\frac{\delta k}{2}\right)\delta^4$$
$$+\mathrm{e}^{\frac{-\delta^2}{2}}\sin\left(\frac{\delta k}{2}\right)\delta^2k^2+2\mathrm{e}^{-\frac{\delta^2}{2}}k\sin\left(\frac{\delta k}{2}\right)\delta^2-2\mathrm{e}^{\frac{-\delta^2}{2}}\sin\left(\frac{\delta k}{2}\right)k^2\Bigg)\Bigg],$$

where $\delta$ is the diameter of the wall. The potential of interaction of the spin electron with the spins of the domain wall $V(x)$ has the form

$$V(x)=J_{sd}S\cos\left(4\arctan\left(\mathrm{e}^{-\delta x}\right)\right). \tag{14}$$

Such potential is plotted in Figure 2. We consider the expansion of Equation (8) in first order. An analysis considering terms of superior order will generate a large quantity of terms in Equation (13) and must not generate any change in the physics properties of the scattering.

We obtain a very complicated expression for the wave function of the electron after the scattering with the ferromagnetic wall domain however, in a combination of two polarization states. The presence of the coefficient $f(k,k')$ in the second term, Equation (3), makes the control of the state of polarization of the electron after the scattering with the domain wall a very difficult analysis.

## 3. Influence of the Spin-Domain Wall Interaction on the Conductivity

The Hamiltonian of the electron that interacts with the domain wall can be written as [22] [23]

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}_{sw}+\mathcal{H}_w, \tag{15}$$

where $\mathcal{H}_0$ is the Hamiltonian of the free electron, $\mathcal{H}_{sw}$ is the electron-domain-wall Hamiltonian and $\mathcal{H}_w$ is the Hamiltonian of the wall domain.

$$\mathcal{H}_0=-t\sum_{\langle ij\rangle}\left(c_{i\uparrow}^{\dagger}c_{j\downarrow}+h.c.\right)+\mu\sum_{i,j}n_{i\uparrow}n_{j\downarrow}, \tag{16}$$

$$\mathcal{H}_{sw}=V\sum_{i,j}S_i\cdot s_j, \tag{17}$$

$$\mathcal{H}_w=J\sum_{i,j}S_i\cdot S_j. \tag{18}$$

$V$ is given by Equation (10).

Making the transformation of the spin operators

$$S_i^+=\sqrt{2S}A_i^{\dagger}, s_i^+=\sqrt{2S}a_i^{\dagger}, \tag{19}$$

$$S_i^-=\sqrt{2S}A_i, s_i^-=\sqrt{2S}a_i, \tag{20}$$

$$S_i^z = S - A_i^\dagger A_i, s_i^z = s - a_i^\dagger a_i, \tag{21}$$

we have for the Hamiltonian $\mathcal{H}_{sw}$

$$\mathcal{H}_{sw} = V \sum_{i,j} \left( A_i^\dagger a_j + h.c. \right) + \mathcal{H}' \tag{22}$$

where $\mathcal{H}'$ contains terms of four or more operators $a_i$ and $A_i$. The contribution of the interaction between electron with domain wall for the electric current operator $\mathcal{J}_{sw}$ is given by

$$\mathcal{J}_{sw} \simeq V \sum_{i,j} \left( A_i^\dagger a_j - h.c. \right). \tag{23}$$

We use the Matsubara's Green function method at finite temperature [22]-[25] to determine the contribution of the interaction of the electron with the wall domain for the regular part of the electric conductivity or continuum conductivity, $\sigma^{reg}(\omega)$ that is given by

$$\sigma_{sw}^{reg}(\omega) = \frac{2V\sqrt{sS}}{\omega} \sum_k \frac{\sin^2 k_x}{\omega_k W_k} \left[ f_k (N_k + 1) \delta(\omega - \omega_k - W_k) + N_k (1 - f_k) \delta(\omega + \omega_k + W_k) \right], \tag{24}$$

as $f_k = \left( e^{\beta \omega_k} + 1 \right)^{-1}$ is the fermion occupation number and $N_k = \left( e^{\beta W_k} - 1 \right)^{-1}$ is the boson occupation number associated with the spin waves of the wall domain and $\beta = 1/T$. We have that in low energy limit

$$\omega_k = v |\boldsymbol{k}|, \tag{25}$$

$$W_k = \frac{J}{3} \left( \cos k_x, + \cos k_y + \cos k_z \right) \tag{26}$$

where $v$ is the Fermi's velocity.

In **Figure 3**, we present the behavior of the contribution of the interaction of the electron with domain wall, $\sigma_{sw}^{reg}(\omega)$. Hence the electric resistance is the inverse of the electric conductivity, the $\sigma_{sw}^{reg}(\omega)$ provides the information about the electric resistance generated by the ferromagnetic domain wall. Our results show a peak of resonance in the contribution spin-electron wall at $\omega \simeq 2.5J$ that indicates a peak in the electric conductivity in this point of $\omega$.

## 4. Conclusions and Final Remarks

In summary, we have studied the scattering between the electrons with the spherical domain wall. We have used the Born's approximation for the $S$ scattering matrix. We obtain a large influence of the scattering on the spin wave function of the electron. We also obtain the contribution of the electron-wall domain interaction on the electric conductivity where it is obtained a peak of resonance in the conductivity for one value of $\omega$ such as $\omega \simeq 2.5J$. We can improve the model Equation (1) with the inclusion of more terms, with objective to get a better description of the scattering of the electron with the wall domain. This is subject to a future work.

From a general way, it is well known that the study of the electron scattering with the

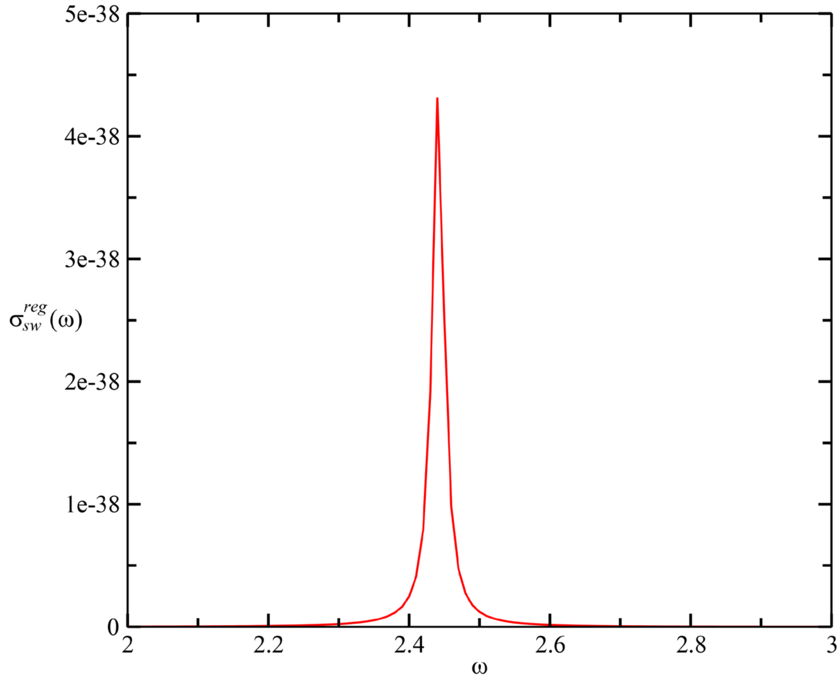



**Figure 3.** Behavior of the contribution of the interaction between electron with the domain-wall, $\sigma_{sw}^{reg}(\omega)$ in the temperature $T$ = 0.1 $J$. The very small value of this contribution is due to interaction of only one electron with the ferromagnetic three-dimensional wall domain. In the electric current we have a flow of $N$ electrons by seconds.

wall domain can generate a different way to generate a spin current spin polarized based on the Hall effect of spin caused by spin-dependent scattering of electrons in thin films [6]. From an experimental point of view, recently there is an intense research about spin transport by electrons where phenomena such as the quantum Hall effect for spins and spintronic [26]-[30] have been studied extensively. In the study of these effects, often only the sign difference between related quantities like magnetic fields can generate the spin and charge currents.

## Acknowledgments

This work was partially supported by the Brazilian agencies FAPEMIG, CAPES, CNPq and CEFET-MG.

## References

[1] Sachdev, S. and Keimer, B. (2011) Physics Today, **64**, 29. http://dx.doi.org/10.1063/1.3554314

[2] Berger, L. (1984) *Journal of Applied Physics*, **55**, 1954. http://dx.doi.org/10.1063/1.333530

[3] Berger, L. (1992) *Journal of Applied Physics*, **71**, 2721. http://dx.doi.org/10.1063/1.351045

[4] Tatara, G. and Kohno, H. (2004) *Physical Review Letters*, **92**, 086601. http://dx.doi.org/10.1103/PhysRevLett.92.086601

[5] Kent, D., Yu, J., Rudiger, U. and Parkin, S. (2001) *Journal of Physics: Condensed Matter*,

**13**, R461. http://dx.doi.org/10.1088/0953-8984/13/25/202

[6] Berger, L. (1986) *Physical Review B*, **33**, 1572. http://dx.doi.org/10.1103/PhysRevB.33.1572

[7] Slonczewski, J.C. (1996) *Journal of Magnetism and Magnetic Materials*, **159**, L1-L7. http://dx.doi.org/10.1016/0304-8853(96)00062-5

[8] Berger, L. (1996) *Physical Review B*, **54**, 9353. http://dx.doi.org/10.1103/PhysRevB.54.9353

[9] Slavin, A. and Tiberkevich, V. (2009) *IEEE Transactions on Magnetics*, **45**, 1875-1918. http://dx.doi.org/10.1109/TMAG.2008.2009935

[10] Sedlmayr, N., Eggert, S. and Sirker, J. (2011) *Physical Review B*, **84**, 024424. http://dx.doi.org/10.1103/PhysRevB.84.024424

[11] Sedlmayr, N. and Berakdar, J. (2012) *Physical Review B*, **86**, 024409. http://dx.doi.org/10.1103/PhysRevB.86.024409

[12] Sedlmayr, N., Ohst, J., Affleck, I., Sirker, J. and Eggert, S. (2012) *Physical Review B*, **86**, 121302(R).

[13] Pereira, R.G. and Miranda, E. (2004) *Physical Review B*, **69**, 140402(R).

[14] Levy, P.M. and Zhang, S.F. (1997) *Physical Review Letters*, **79**, 5110. http://dx.doi.org/10.1103/PhysRevLett.79.5110

[15] Baibich, M.N., Broto, J.M., Fert, A., Nguyen Van Dau, F., Petroff, F., Eitenne, P., Creuzet, G., Friederich, A. and Chazelas, J. (1988) *Physical Review Letters*, **61**, 2472. http://dx.doi.org/10.1103/PhysRevLett.61.2472

[16] Abdel-Aty, A.-H., Zakaria, N., Cheong, L.Y. and Metwally, N. (2014) *Journal of Quantum Information Science*, **4**, 1-17. http://dx.doi.org/10.4236/jqis.2014.41001

[17] Rippard, W.H., Pufall, M.R., Kaka, S., Russek, S.E. and Silva, T.J. (2004) *Physical Review Letters*, **92**, Article ID: 027201. http://dx.doi.org/10.1103/PhysRevLett.92.027201

[18] Krivorotov, L.N., Emley, N.C., Sankey, J.C., Kiselev, S.I., Ralph, D.C. and Buhrman, R.A. (2005) *Science*, **307**, 228-231. http://dx.doi.org/10.1126/science.1105722

[19] Rezende, S.M., de Aguiar, F.M. and Azevedo, A. (2005) *Physical Review Letters*, **94**, Article ID: 037202. http://dx.doi.org/10.1103/PhysRevLett.94.037202

[20] Pesin, D. and MacDonald, A.H. (2012) *Nature Materials*, **11**, 409-416. http://dx.doi.org/10.1038/nmat3305

[21] Chumak, A.V., Vasyuchka, V.I., Serga, A.A. and Hillebrands, B. (2015) *Nature Physics*, **11**, 453-461. http://dx.doi.org/10.1038/nphys3347

[22] Fradkin, E. (2013) Field Theories of Condensed Matter Physics. 2nd Edition, Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9781139015509

[23] Wen, X.-G. (2010) Quantum Field Theory of Many-Body Systems. Oxford Graduate Texts, New York.

[24] Kubo, R., Toda, M. and Hashitsume, N. (1985) Statistical Physics II. Springer-Verlag, New York. http://dx.doi.org/10.1007/978-3-642-96701-6

[25] Mahan, G.D. (1990) Many Particles Physics. Plenum, New York. http://dx.doi.org/10.1007/978-1-4613-1469-1

[26] Lauer, V., Bozhko, D.A., Bracher, T., Pirro, P., Vasyuchka, V.I., Serga, A.A., Jungfleisch, M.B., Agrawal, M., Kobljanskyj, Y.V., Melkov, G.A., Dubs, C., Hillebrands, B. and Chumak, A.V. (2015.) arXiv:1508.07517v1

[27] Meier, F. and Loss, D. (2003) *Physical Review Letters*, **90**, Article ID: 167204.

[28] Kobljanskyj, Y.V., Melkov, G.A., Serga, A.A., Slavin, A.N. and Hillebrands, B. (2015) *Phys-*

*ical Review Applied*, **4**, Article ID: 014014.
http://dx.doi.org/10.1103/PhysRevApplied.4.014014

[29] Han, W., McCreary, K.M., Pi, K., Wang, W.H., Li, Y., Wen, H., Chen, J.R. and Kawakami, R.K. (2012) *Journal of Magnetism and Magnetic Materials*, **324**, 369-381.
http://dx.doi.org/10.1016/j.jmmm.2011.08.001

[30] Azevedo, A., Vilela-Leao, L.H., Rodriguez-Suarez, R.L., Lacerda Santos, A.F. and Rezende, S.M. (2011) *Physical Review B*, **83**, Article ID: 144402.
http://dx.doi.org/10.1103/PhysRevB.83.144402

Journal of Modern Physics, 2016, 7, 1644-1650
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Effects of Band Nonparabolicity and Band Offset on the Electron Gas Properties in *InAs/AlSb* Quantum Well

**Gafur Gulyamov[1], Bahrom Toshmirza O'g'li Abdulazizov[2], Baymatov Paziljon Jamoldinovich[2]**

[1]Namangan Engineering-Pedagogical Institute, Namangan, Uzbekistan
[2]Namangan State University, Namangan, Uzbekistan
Email: Gulyamov1949@mail.ru

**How to cite this paper:** Gulyamov, G., Abdulazizov, B.T. and Jamoldinovich, B.P. (2016) Effects of Band Nonparabolicity and Band Offset on the Electron Gas Properties in *InAs/ AlSb* Quantum Well. *Journal of Modern Physics*, **7**, 1644-1650.
http://dx.doi.org/10.4236/jmp.2016.713149

**Received:** July 13, 2016
**Accepted:** September 10, 2016
**Published:** September 13, 2016



## Abstract

One-band effective mass model is used to simulation of electron gas properties in quantum well. We calculate of dispersion curves for first three subbands. Calculation results of Fermi energy, effective mass at Fermi level as function of electron concentration are presented. The obtained results are good agreement with the experimental dates.

## Keywords

Quantum Well, In-Plane Dispersion, *InAs*, *AlSb*, Two Dimentional Electron Gas, Effective Mass, Cyclotron Mass

## 1. Introduction

In semiconductors, *InAs* and *InSb* of the conduction band are characterized by a strong nonparabolicity and recently intensively studied heterostructures based on them [1]-[3]. Nonparabolicity of the conduction band and the nature of the spin splitting of the electron in the quantum well (QW) are studied by the cyclotron resonance [4]-[7].

In [8] has been investigated *InAs/AlSb* based QW with well width $L = 15$ nm, where two dimensional (2D) electron concentration ranges from $2.7 \times 10^{11}$ to $8 \times 10^{12}$ cm$^{-2}$. In this work has been found increase of the effective mass of almost 2 times.

The purpose of this work—the calculation of: 1) subbands dispersion curves, 2) the density of states of 2D electron gas and 3) concentration dependence of effective mass in Fermi level for *InAs/AlAs* QW with width $L = 15$ nm.

It is shown that an abrupt change in the density of states leads to a peculiar change in the concentration dependence of effective mass.

## 2. The In-Plane Dispersion

Consider a single QW with width $L$ (area *A*—*InAs*), concluded between barriers with height $V$ (area *B*—*AlAs*). The energy is measured from the bottom of the band of the bulk *InAs*.

In the one band effective mass approximation, the solution of the three-dimensional Schrödinger equation can be represented as $\psi = \mathrm{e}^{i(k_x x + k_y y)}\phi(z)$. Then for the area $A$ and $B$, respectively, we can write the following one-dimensional equations

$$\frac{\partial^2 \phi_A(z)}{\partial z^2} + q^2 \phi_A(z) = 0, \quad q = \sqrt{\frac{2m_A(E)}{\hbar^2} E - k^2} \tag{1}$$

$$\frac{\partial^2 \phi_B(z)}{\partial z^2} + \chi^2 \phi_B(z) = 0, \quad \chi = \sqrt{\frac{2m_B(E)}{\hbar^2}(V - E) + k^2} \tag{2}$$

Here $k^2 = k_x^2 + k_y^2$ and $k-$ is in plane wave vector, $m_{A,B}(E)$ —energy-dependent effective mass of the electrons in the material $A$ or $B$. Solving Equations (1) and (2), using the boundary condition

$$\phi_A(0) = \phi_B(0), \quad \frac{1}{m_A(E)} \frac{\mathrm{d}\phi_A(z)}{\mathrm{d}z}\bigg|_{z=0} = \frac{1}{m_B(E)} \frac{\mathrm{d}\phi_B(z)}{\mathrm{d}z}\bigg|_{z=0} \tag{3}$$

$$\phi_A(L) = \phi_B(L), \quad \frac{1}{m_A(E)} \frac{\mathrm{d}\phi_A(z)}{\mathrm{d}z}\bigg|_{z=L} = \frac{1}{m_B(E)} \frac{\mathrm{d}\phi_B(z)}{\mathrm{d}z}\bigg|_{z=L} \tag{4}$$

we find the dispersion equation

$$E = E_\parallel + E_0 \left[ \pi \cdot n - 2\, Arc\sin\left( \sqrt{\frac{\gamma^2 (E - E_\parallel)}{\gamma(\gamma - 1)E + \gamma\, V + (1 - \gamma^2)E_\parallel}} \right) \right]^2 \tag{5}$$

here $E_\parallel = \hbar^2 k^2 / 2m_A(E)$, $E_0 = \hbar^2 / 2m_A(E) L^2$, $\gamma = m_B(E)/m_A(E)$.

Nonparabolicity of conduction band well takes into account by formulas

$$\frac{m_0}{m_A(E)} = 1 + \frac{E_{PA}}{3}\left[ \frac{2}{E + E_{gA}} + \frac{1}{E + E_{gA} + \Delta_A} \right] \tag{6}$$

$$\frac{m_0}{m_B(E)} = 1 + \frac{E_{PB}}{3}\left[ \frac{2}{E - V + E_{gB}} + \frac{1}{E - V + E_{gB} + \Delta_B} \right] \tag{7}$$

where, $m_0$ —the free electron mass, $E_P$ —the Kane parameter, $E_g$ —the band gap, $\Delta$ — the spin-orbital splitting of valence band, $V$—conduction band offset. Band parameters of *InAs* and *AlSb* are shown in Table 1.

To describe the statistics of electrons, Equation (5) is non convenient because it is not solvable with respect to $E$ or $k$. Therefore, we replace Equation (5) is by simple approximation

$$E \frac{m_A(E)}{m_A(0)} \approx \frac{\hbar^2 k^2}{2m_A(0)} + E_n \frac{m_A(E_n)}{m_A(0)} \tag{8}$$

**Table 1.** Band parameters of *InAs* and *AlSb*.

| | *InAs* (*A*) | *AlSb* (*B*) |
|---|---|---|
| $E_g$, eV | 0.42 | 2.37 |
| $\Delta$, eV | 0.38 | 0.75 |
| $E_P$, eV | 21.2 | 20.85 |
| $m(0)$, $[m_0]$ | 0.023 | 0.11 |
| $V$, eV | 0 | 1.35 |

where, $E_n$—is bottom of n-th subbands. Now, approximation (8) is the best solution of (5). However, values of $E_n$ in (8) now are obtained from Equation (5) at $k = 0$ by use numeric method.

For *InAs/AlSb* QW with $L = 15$ nm, we have: $E_1 = 0.0454$ eV, $E_2 = 0.158$ eV, $E_3 = 0.304$ eV, and for case $L = 6$ nm we have: $E_1 = 0.163$ eV, $E_2 = 0.509$ eV, $E_3 = 0.903$ eV.

Calculated dispersion curves from Equation (5) and approximation (8) are compared in **Figure 1(a)**, **Figure 1(b)**.

From **Figure 1(a)**, **Figure 1(b)** follows that, the approximation (8) is sufficiently accurate and/or (8) is the best solution of (5) in a wide range of width QW. It is convenient when studying the statistics of electrons, kinetic, optical, or other characteristics of the 2D electron gas. Inconveniences approximation (8) is such that $E_n\left(n=1,2,3,\cdots\right)$ depends on $L$, $V$ and on other parameters of materials $A$,$B$. Therefore every time when changing these parameters, the value of $E_n$ is recalculated from Equation (5) at $k = 0$.

## 3. The Fermi Energy and Thermodynamic DOS

The total electron concentration is

$$n_s = N_0 \sum_{n=1} \int_{E_n}^{E_F} \mathrm{d}\left(\frac{\hbar^2 k^2}{2m_A(0)}\right) \tag{9}$$

where

$$N_0 = \frac{m_A(0)}{\pi\hbar^2} = \frac{m_A(0)\cdot 0.413\times 10^{15}}{\text{эВ}\cdot\text{см}^2}.$$

According (8) we have

$$n_S = \sum_{n=1} n_S^{(n)} \quad , \quad n_s^{(n)} = N_0\left[E_F \frac{m_A(E_F)}{m_A(0)} - E_n \frac{m_A(E_n)}{m_A(0)}\right]\theta\left(E_F - E_n\right) \tag{10}$$

where $n_s^{(n)}$—is $n$-th subband concentration.

In Equation (10), the terms in the sum should be positive. The negative terms in $n_s^{(n)}$ excluded by the Heaviside function. It establishes a link between the Fermi energy $E_F$ and full 2D electron concentration $n_s$. They also determine the concentration of electrons in separate subbands $n_s^{(n)}$ for a given $n_s$.

From (10), we can estimate the critical concentrations of $n_{c1}$, in which the Fermi level comes to bottom of the second subband $E_F = E_2$. In the structure of InAs/AlSb QW

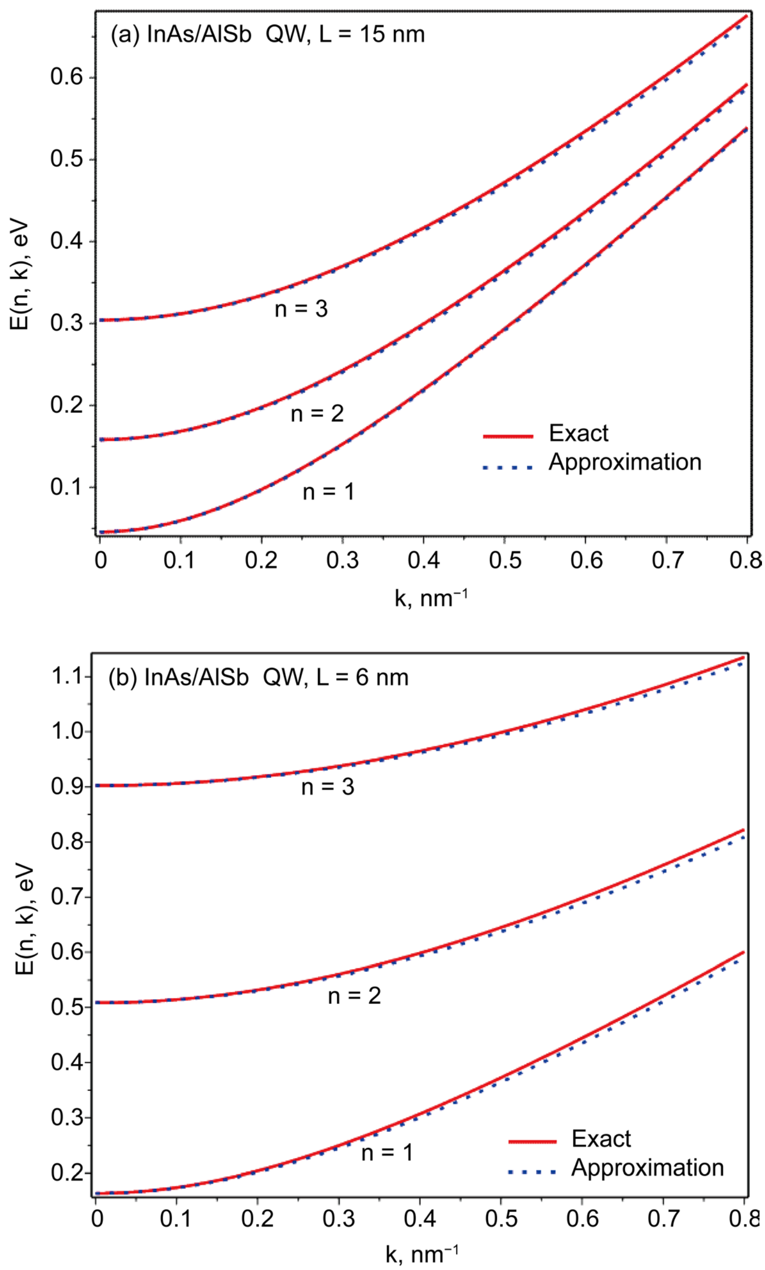



**Figure 1.** The dispersion curves of the first three subbands ($n = 1, 2, 3$) in *InAs/AlSb* QW: the full (red) line—according to Equation (5), the dotted (blue) line—approximation (8); (a) $L = 15$ nm, (b) $L = 6$ nm.

with the well width $L = 15$ nm can be found

$$n_{c1} = N_0 \left[ E_2 \frac{m_A(E_2)}{m_A(0)} - E_1 \frac{m_A(E_1)}{m_A(0)} \right] = 1.52 \times 10^{12} \text{ cm}^{-2} \tag{11}$$

This estimation is close to experimental measured date $n_{c1,\exp} = 1.2 \times 10^{12}$ cm$^{-2}$ [8]. Similarly, the critical concentration of $n_{c2}$, in which the Fermi level comes to bottom of the third subband $E_F = E_3$, is $n_{c2} = 6.87 \times 10^{12}$ cm$^{-2}$.

The dependence $E_F(n_s)$ is shown in **Figure 2**. It is obtained from Equations (10) by changing the Fermi energy in the range of $E_F = 0 \div 0.33\ \text{eV}$. The graph shows that, depending on $E_F(n_s)$ there exist a fractures—slowing of increase the Fermi's energy. They are caused by abrupt changes (by jumps), the density of states at the critical points: $n_s = n_{c1}$, $E_F = E_2$, $n_s = n_{c2}$, $E_F = E_3$.

These fractures occur at the critical concentrations of $n_{c1}, n_{c2} \cdots$, where $E_F$ intersects the Fermi level of the bottom of the next subband.

The thermodynamically DOS of electron gas at Fermi level $N_T(E_F) = \mathrm{d}n_S/\mathrm{d}E_F$ is shown in **Figure 3**.

## 4. The Cyclotron Mass

According approximation (8), the electron effective mass at the Fermi level (cyclotron mass) $m_c = \hbar^2 k \left(\mathrm{d}E/\mathrm{d}k\right)^{-1}\Big|_{E=E_F}$ is

$$m_c = \left( m_A(E) + E\frac{\mathrm{d}m_A(E)}{\mathrm{d}E} \right)_{E=E_F} \quad (12)$$

The dependence $m_c(n_s)$ is shown in **Figure 4**.

This dependence can be obtained from Equations (10) and (12) by changing the Fermi energy in the range $E_F = 0 - 0.33\ \text{eV}$.

This figure shows also the dependence of experimentally measured value of the effective masses (cyclotron mass) $m_c$ at the Fermi level of the total concentration of 2D $n_s$ [8]. Fracture in the $m_c(n_s)$ occur at critical points: $n_s = n_{c1}$, $E_F = E_2$ and $n_s = n_{c2}$, $E_F = E_3$.

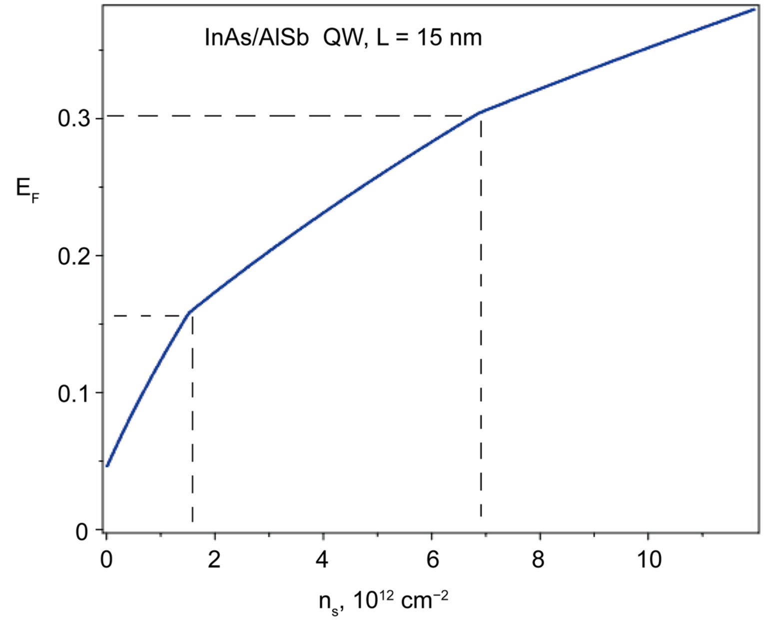



**Figure 2.** The dependence of the Fermi energy on the 2D concentration in *InAs/AlSb* QW with $L$ = 15 nm, $V$ = 1.35 eV. The critical points $n_s = n_{c1}$, $E_F = E_2$ and $n_s = n_{c2}$, $E_F = E_3$ shown by the shaded lines.

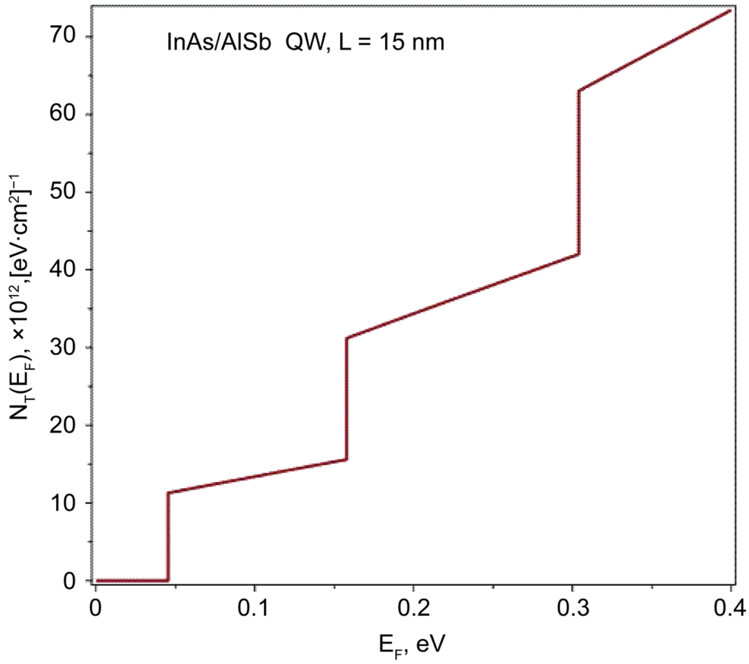



**Figure 3.** The thermodynamically DOS of electron gas at Fermi level in *InAs/AlSb* QW, $L$ = 15 nm, $V$ = 1.35 eV.

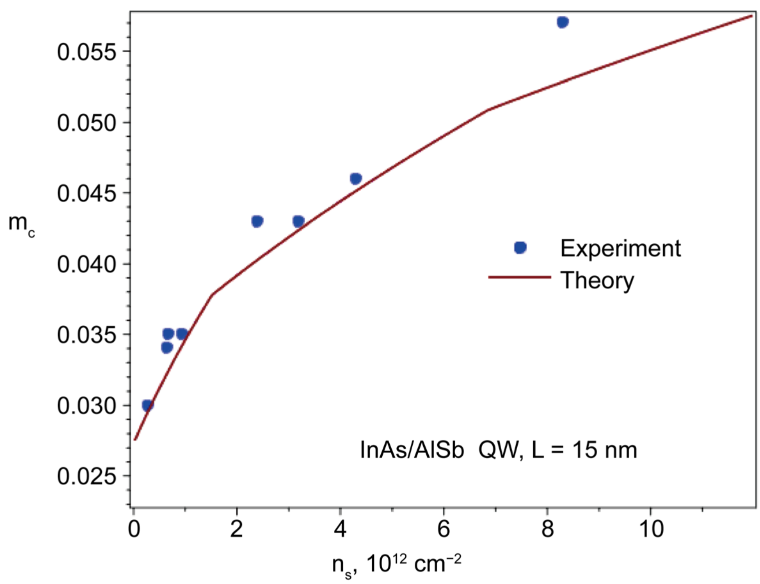



**Figure 4.** The dependence of the effective mass on the total concentration for InAs/AlSb QW, $L$ = 15 нм, $V$ = 1.35 eV.

## 5. Conclusion

In this study are provided useful approximation (8) of subband dispersions and simplified Equation (10) to calculate the statistics of a degenerate electron gas in heterostructured InAs/AlSb QW, which satisfactorily describes the experimental results [8]. They are also useful to study of calculation of the transport, optical and magnetic properties of electron gas in a Kane type 2D system. The above description of the algorithms can be applied to other QW heterostructures based on semiconductor group $A_3B_5$.

## Acknowledgements

This work was supported by the Scientific and Technical program Republic of of Uzbekistan (Grant F2-OT-O-15494).

## References

[1] Vasilyev, Yu.B., Gouider, F., Nachtwei, G. and Buckle, P.D. (2010) *Semiconductors*, **44**, 1511-1514. http://dx.doi.org/10.1134/S1063782610110266

[2] Spirin, K.E., Kalinin, K.P., Krishtopenko, S.S., Maremyanin, K.V., Gavrilenko, V.I. and Sadofyev, Yu.G. (2012) *Semiconductors*, **46**, 1396-1401. http://dx.doi.org/10.1134/S1063782612110206

[3] Barate, D., Teissier, R., Wang, Y. and Baranov, A.N. (2005) *Applied Physics Letters*, **87**, 051103. http://dx.doi.org/10.1063/1.2007854

[4] Yang, M.J., Lin-Chung, P.J., Wagner, R.J., Waterman, J.R., Moore, W.J. and Shanabrook, B.V. (1993) *Semiconductor Science and Technology*, **8**, S129. http://dx.doi.org/10.1088/0268-1242/8/1S/029

[5] Yang, M.J., Lin-Chung, P.J., Shanabrook, B.V., Waterman, J.R., Wagner, R.J. and Moore, W.J. (1993) *Physical Review B*, **47**, 1691. http://dx.doi.org/10.1103/PhysRevB.47.1691

[6] Gauer, C., Scriba, J., Wixforth, A., Kotthaus, J.P., Bolognesi, C.R., Nguyen, C., Brar, B. and Kroemer, H. (1994) *Semiconductor Science and Technology*, **9**, 1580. http://dx.doi.org/10.1088/0268-1242/9/9/002

[7] Warburton, R.J., Gauer, C., Wixforth, A., Kotthaus, J.P., Brar, B. and Kroemer, H. (1996) *Physical Review B*, **53**, 7903. http://dx.doi.org/10.1103/PhysRevB.53.7903

[8] Aleshkin, V.Ya., Gavrilenko, V.I., Ikonnikov, A.V., Sadofyev, Yu.G., Bird, J.P., Jonhson, S.R. and Zhang, Y.-H. (2005) *Semiconductors*, **39**, 62-66. http://dx.doi.org/10.1134/1.1852647

Journal of Modern Physics, 2016, 7, 1651-1660
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Counterfactual Definiteness and Bell's Inequality

**Karl Hess[1], Hans De Raedt[2], Kristel Michielsen[3,4]**

[1]Center for Advanced Study, University of Illinois, Urbana, IL, USA
[2]Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands
[3]Institute for Advanced Simulation, Jülich Supercomputing Centre, Jülich, Germany
[4]RWTH Aachen University, Aachen, Germany
Email: k.michielsen@fz-juelich.de

**How to cite this paper:** Hess, K., De Raedt, H. and Michielsen, K. (2016) Counterfactual Definiteness and Bell's Inequality. *Journal of Modern Physics*, **7**, 1651-1660.
http://dx.doi.org/10.4236/jmp.2016.713150

**Received:** May 23, 2016
**Accepted:** September 11, 2016
**Published:** September 14, 2016



## Abstract

Counterfactual definiteness must be used as at least one of the postulates or axioms that are necessary to derive Bell-type inequalities. It is considered by many to be a postulate that not only is commensurate with classical physics (as for example Einstein's special relativity), but also separates and distinguishes classical physics from quantum mechanics. It is the purpose of this paper to show that Bell's choice of mathematical functions and independent variables implicitly includes counterfactual definiteness. However, his particular choice of variables reduces the generality of his theory, as well as the physics of all Bell-type theories, so significantly that no meaningful comparison of these theories with actual Einstein-Podolsky-Rosen experiments can be made.

## Keywords

Foundations of Quantum Mechanics, Foundations of Probability, Bell Inequality

## 1. Introduction

Bell's theorem [1] has an unusual standing among mathematical-physical theorems. No other theorem has ever been discussed with respect to so many "loopholes", physical situations that make it possible to escape the mathematical strictures of the theorem. It is shown that the reason for this fact is that Bell's theorem is based on the postulate of counterfactual definiteness. The postulate of counterfactual definiteness to derive Bell-type inequalities is clearly asserted in the books of Peres [2] and Leggett [3].

Some of Einstein's reasoning regarding Einstein-Podolsky-Rosen (EPR) experiments also contain counterfactual realism and Einstein's special relativity is counterfactually definite in the mathematical sense presented below. This fact may have contributed to the opinion that counterfactual realism is the major defining trait of "classical" theories. It will be shown, however, that great care must be exercised with respect to the choice of independent variables in the argument of the functions that are used to formulate a counterfactually definite physical theory. It will also be shown that the particular choice of variables, used for the derivation of Bell's inequality and Bell's theorem, imposes significant restrictions to the physical situations that can be described by Bell's functions and excludes dynamic processes of classical physics, no matter whether deterministic or stochastic. To show this fact, we first repeat the main features of Bell's functions that describe Einstein-Podolsky-Rosen-Bohm (EPRB) experiments and then connect them to a precise definition of counterfactual definiteness.

## 2. EPRB Experiments and Bell's Functions Representing Them

EPRB experiments have been described extensively in the literature [4] [5] and measure the spins of entangled particle pairs at two space-like separated locations. The two particles of each pair are emanating from a source and propagate toward the space-like separated locations. The properties of these particles are measured by instruments that are described by a "setting" such as the direction of a polarizer or magnet which is characterized by a unit vector of three dimensional space denoted by $\boldsymbol{j} = \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \cdots$. Measurements of this type have been performed by a number of researchers and have had a checkered history with respect to the results. These, at first, contradicted and then confirmed quantum theory [4]. There are still significant deviations from quantum theory in current experiments, which are, however, mostly ignored [6]. We proceed here by just stipulating that indeed these experiments showed a violation of the, by now, famous Bell inequality and describe in the following only Bell's postulates and assumptions, thereby focusing on the simplest case involving only three settings and not four, as used in actual experiments, see also [7]. Bell's postulates and assumptions are considered by many researchers to be entirely general and valid for all EPR like experiments and Gedanken-experiments as long as they can be described by classical physics such as Einstein's relativity.

Bell's classical-physics model for the system of measurement equipment and entangled pairs of the EPRB experiments is constructed as follows (see page 8 of [1]). He assumed that all experimental results, all data, can be described by using functions $A$ that map the independent measurement results onto $\pm 1$ which symbolizes the two possible outcomes of the spin measurements. The variables in the argument of the function always include the settings $\boldsymbol{j} = \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \cdots$ and another variable, or set of variables, that Bell denoted by $\lambda$. Bell then proceeded to present a proof of his now celebrated inequality:

$$\left\langle A(\boldsymbol{a},\lambda)A(\boldsymbol{b},\lambda)+A(\boldsymbol{a},\lambda)A(\boldsymbol{c},\lambda)-A(\boldsymbol{b},\lambda)A(\boldsymbol{c},\lambda)\right\rangle \leq +1, \tag{1}$$

where $\langle\cdot\rangle$ indicates the average over many measurements. The left and right factor of each term correspond to the data taken at the two corresponding space like separated measurement stations. The events of measurements and corresponding data are linked to clock times of two synchronized laboratory clocks. Therefore, the functions $A$ as well as the variables $\boldsymbol{j}$ and $\lambda$ must, for each of the products, correspond to pairs of clock times $t_n, t_{n'}$ where $n$ is the measurement number. These clock times are not explicitly included as indexes or variables of Bell's functions.

Note that Bell's original paper (see page 7 of [1]) assigned to $\lambda$ only properties of the entangled pair. It is now generally assumed [8] that $\lambda$ may stand for a set of arbitrary physical variables including space and time coordinates or even Einstein's space-time ***st***. Therefore, $\lambda$ may also describe some properties of the measurement equipment (in addition to the magnet or polarizer orientation $\boldsymbol{j}$), such as dynamical effects arising from many-body interactions of the entangled pair with the constituent particles and fields of the measurement equipment. Bell agreed with this assumption in his later work [1].

It is the purpose of this paper to show that the postulate of counterfactual definiteness in conjunction with the use of a setting variable $\boldsymbol{j}$ does not permit the introduction of general space and time related variables that describe the said many body dynamics. Therefore, Bell's assumptions are not general enough to describe classical theories of EPRB experiments that include dynamic processes involving the measurement equipment.

## 3. Counterfactual Reasoning and EPRB Experiments

Peres [2] gave the following definition of counterfactual realism, which roughly agrees with the definition of Leggett [3]. Peres claims, as does Leggett, not to use traditional concepts of mathematics and physics to start with, but only "what could have possibly been the results of unperformed experiments" and bases his definition of counterfactual realism on the following statement:

> It is possible to imagine hypothetical results for any unperformed test, and to do calculations where these unknown results are treated as if they were numbers.

We agree that it is possible, as a purely intellectual activity, to imagine hypothetical results for any unperformed tests. However, without significant additional assumptions, it is not possible "to do calculations where these unknown results are treated as if they were numbers". Here we encounter the so often unrecognized gulf between sense impressions, even just imagined ones, and conceptual frame-works such as the axiomatic system of numbers or the probability theory of Kolmogorov. Peres, Leggett and a majority of quantum information theorists did not and do not recognize that giant gulf, that giant separation, between events of nature, recorded as data, and the axiomatic edifices of human thought.

If Peres wishes to treat hypothetical "results" of unperformed tests as if they were numbers, he must be sure that these abstractions at least follow the axioms of numbers. There are several steps necessary to connect the "events" of the physical world to numbers. Boole derived ultimate alternatives and a Boolean algebra while Kolmogorov's axiomatic system introduces an event algebra and probability space. It is true that mathematicians often describe experimental situations or ideas about them by the Kolmogorov framework and just postulate that a probability space and $\sigma$-algebra exists. It is known, however, from the work of Boole [9] and Vorob'ev [10] that a given particular set of variables may not be able to describe certain correlations in any given set of data.

In more elementary terms, we have to consider the following facts. If we perform "calculations where these unknown results are treated as if they were numbers", then we must use the mathematical concept of functions or something equivalent in order to link the imagined but possible tests with numbers. A one to one correspondence of the possible tests and the numbers needs to be established and it needs to be shown that no logical-mathematical contradictions arise from such procedure. If no such correspondence exists, then the "purely intellectual activity" is nothing more than child's play and the mathematical abstractions of such activity can certainly not be treated as if they were numbers with some relation to physics.

Take any set of data derived from measurements on spin-1/2 particles with Stern-Gerlach magnets, that lists the measured spins as "up" or "down" together with magnet settings $\boldsymbol{j} = \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \cdots$. Can we replace "up" with +1 and "down" with −1 and expect that the so obtained set follows the axioms of integers? The "trespass" to deal with tests as if they were numbers has been committed by several textbook authors, in particular by Peres [2] and Leggett [3]. This point appears in clear relief, if we write down the data according to the way in which they are imagined to be taken in testing e.g. the Bell-type inequality. The data are recorded in pairs corresponding to detector-events that are registered together with equipment settings and the clock times of synchronized laboratory clocks. Thus we obtain data lists of the kind: $\left(D_{j_1}^{t_1}, D_{j'_1}^{t'_1}\right), \left(D_{j_2}^{t_2}, D_{j'_2}^{t'_2}\right), \cdots, \left(D_{j_N}^{t_M}, D_{j'_M}^{t'_M}\right)$, the $\boldsymbol{j}_n, \boldsymbol{j}_{n'}$ representing the randomly chosen setting pair and $t_n, t'_n$ denoting the times of measurement. Here the $D$'s are symbols that represent the measured up/down spin in the example above but may as well represent the red/green color of a flash of light, etc. For numerical processing of this list of symbols, it is expedient to introduce new symbols $\hat{D}_{\boldsymbol{j}}^{t}$ taking values +1 and −1 that are in one-to-one correspondence with the original symbols $D_{\boldsymbol{j}}^{t}$. The number of times that the setting ($\boldsymbol{a}$, $\boldsymbol{b}$), ($\boldsymbol{a}$, $\boldsymbol{c}$), and ($\boldsymbol{b}$, $\boldsymbol{c}$) was chosen is denoted by $N_{\boldsymbol{a},\boldsymbol{b}}$, $N_{\boldsymbol{a},\boldsymbol{c}}$, and $N_{\boldsymbol{b},\boldsymbol{c}}$, respectively. The total number of pairs is then $M = N_{\boldsymbol{a},\boldsymbol{b}} + N_{\boldsymbol{a},\boldsymbol{c}} + N_{\boldsymbol{b},\boldsymbol{c}}$. One cannot do justice to the number of different data-pairs by using models with three pairs of mathematical symbols such as $A_{\boldsymbol{a}}$, $A_{\boldsymbol{b}}$, $A_{\boldsymbol{a}}$, $A_{\boldsymbol{c}}$, and $A_{\boldsymbol{b}}$, $A_{\boldsymbol{c}}$ as they are used in Bell-type proofs. One runs into problems even if one regards these mathematical symbols as "variables" (such as Boolean variables [11]) and not just as numbers; the reason being that one cannot cover all the different possible correlations of the data by such few variables. If we admit the two values +1 and −1 for the $\hat{D}$'s at different times of the same experiment, then we obtain $N_{\boldsymbol{a},\boldsymbol{b}} + 1$ different values for the sum of the pair product $\sum_{n=1}^{M} \delta_{j_n,\boldsymbol{a}} \delta_{j'_n,\boldsymbol{b}} \hat{D}_{\boldsymbol{a}}^{t_n} \hat{D}_{\boldsymbol{b}}^{t'_n}$. If we have three such sums with all independent variables, the number of possibilities is $\left(N_{\boldsymbol{a},\boldsymbol{b}}+1\right)\left(N_{\boldsymbol{a},\boldsymbol{c}}+1\right)\left(N_{\boldsymbol{b},\boldsymbol{c}}+1\right) \approx \left(M/3+1\right)^3$ for $M$ sufficiently large. In contrast, we have for the Bell type variables $A_{\boldsymbol{a}}$, $A_{\boldsymbol{b}}$, $A_{\boldsymbol{a}}$, $A_{\boldsymbol{c}}$, and $A_{\boldsymbol{b}}$, $A_{\boldsymbol{c}}$ only about $\left(M/3+1\right)^2$ independent choices of all possible different correlations of possible outcomes of these variables. This fact arises from Bell's description of 3 $M$ different **pairs** of measurements (6 $M$ measurements) by only 3 different variables and represents another typical trespass that is explicitly made in both the book of Peres [2] and Leggett [3]: they use a model with a severe restriction of choices before any physics is introduced and thus"overburden" their variables in a way which cannot do justice to the complexity of the data. In real EPRB experiments, one uses four not three different randomly chosen settings [12] [13] but the above argument equally holds for this case, with $(M/3+1)^3$ and $(M/3+1)^2$ being replaced by $(M/4+1)^4$ and $(M/4+1)^3$ for 4 $M$ different pairs (8 $M$ measurements), respectively.

This more subtle problem, a well known problem in the area of computer simulations, reveals once more the enormous gulf between data and mathematical abstractions that describe the data. In the framework of Boole [11], we need to be sure that the data can be described by ultimate alternatives (the Boolean variables) and in the framework of Kolmogorov we must be sure to deal with random variables (functions on a Kolmogorov probability space). But how can we be sure? As a minimum requirement we need to introduce functions, with sufficiently many physical variables in their arguments, to enable the description of all the possible correlations and to guarantee a one to one correspondence of mathematical abstractions and the massive amount of data.

To describe EPRB experiments in the general way that Bell intended and purported to actually have done, we need to introduce functions $A$ with variables additional to $\boldsymbol{j}$ in their argument (or indexes, see below). We need to have variables such as $t_n$, $s_n$, $\boldsymbol{st}_n$, ⋯ that are taken out of the realm of Einsteinian physics and do indeed guarantee the one to one correspondence to the data. For example, we may need to include $t_n$, the time of measurement at one location and $s_n$ representing any property of the objects emanating from the source. It may also be necessary to include a more general four dimensional space-time vector $\boldsymbol{st}_n$ instead or in addition to the measurement time $t_n$ and we include it here just for completeness. This way we obtain functions $A = A(\boldsymbol{j}_n, t_n, s_n, \boldsymbol{st}_n, \cdots)$.

Some may ask whether that is not precisely what Bell used by introducing his $\lambda$ that, as he claimed [1], can stand for any set of variables and, therefore, also for the set $(t_n, s_n, \boldsymbol{st}_n, \cdots)$. We thus may have $A = A(\boldsymbol{j}_n, t_n, s_n, \boldsymbol{st}_n, \cdots) = A(\boldsymbol{j}_n, \lambda_n)$. Indeed it is true that this is what Bell claimed. However, as we will see below his claim is incorrect, because he and followers have postulated complete independence of $\lambda$ and $\boldsymbol{j}$ and thus postulated counterfactual definiteness in conjunction with the setting variable $\boldsymbol{j}$ according to the precise definition given in the next section. Einstein locality does not require independence of $\lambda$ of the local setting (see Section 5).

Note that quantum mechanics does not use any setting-type of variable as independent variable in the argument of the wave-function. There, the setting-type variables label the operators. A helpful discussion of explicit and implicit assumptions of Bell, with emphasis of the mathematical structure and consistency, was given by Khrennikov [14].

## 4. Mathematical Definition of Counterfactual Definiteness and Bell's Inequality

Counterfactual definiteness requires the following. We must be able to describe a measurement or test by using a given set of variables in the argument of the function $A$, and thus for example a setting $\boldsymbol{j} = \boldsymbol{b}$. Then, we must also be able to reason that we could have used instead of setting $\boldsymbol{b}$ the setting $\boldsymbol{c}$ and would have obtained the outcome corresponding to the value of $A$, now calculated with setting $\boldsymbol{c}$ and all other variables in its argument unchanged. Although this type of reasoning is not permitted in the courts of law, its mathematical restatement looks natural and general enough:

> *A counterfactually definite theory is described by a function (or functions) that map(s) tests onto numbers. The variables of the function(s) argument(s) must be chosen in a one to one correspondence to physical entities that describe the test(s) and must be independent variables in the sense that they can be arbitrarily chosen from their respective domains.*

This definition means that the outcomes of measurements must be described by functions of a set of independent variables. The definition applies, of course, to the major theories of classical physics, including Einstein's special relativity. Counterfactual definiteness appears, therefore, as a reasonable and even necessary requirement of classical theories. However, most importantly, counterfactual definiteness restricts the use of variables to those that can be independently picked from their respective domains. However, a magnet- or polarizer-orientation, mathematically represented by the variable $\boldsymbol{j}$, cannot be picked independently of the measurement times, which are mathematically represented by $t_n$ and registered by the clocks of the measurement stations. Once a setting is picked at a certain space-time coordinate, no other setting can be linked to that coordinate, because of the relativistic limitations for the movement of massive bodies and the fact that Bell's theory is confined to the realm of Einsteinian physics and, therefore, excludes quantum superpositions. Thus any measurement is related to spatio-temporal equipment changes and the mathematical variables that describe the measurement need to represent the possible physical situations.

Enter probability theory and we certainly cannot use the setting $\boldsymbol{j}$ as a random variable and the measurement time $\boldsymbol{t}$ as another *independent* random variable on the same probability space. The reason for this fact is rooted in the above explanation and can be further crystallized as follows. It is possible to define the setting $\boldsymbol{j}$ as a random variable on one probability space meaning that we may regard $\boldsymbol{j}$ as a function which assigns to each elementary event $\omega$ of a sample space $\Omega$ a so called realization of $\boldsymbol{j}$ e.g. $\boldsymbol{j}(\omega_1) = \boldsymbol{b}$. It is also possible, at least under very general circumstances, to formulate the measurement times as another random variable $t(\omega')$, where $\omega'$ is an elementary event of a second sample space $\Omega'$. Again, given some specific $\omega_1'$ we obtain a realization e.g. $t(\omega_{1'}) = t_1$.

However, the formation of a product probability space on which both random variables $\boldsymbol{j}$ and $t$ are defined presents now a problem. That space would necessarily contain impossible events (such as different settings for the same measurement times) with a non-zero product probability measure assigned to them. These facts can actually be formulated as a theorem stating that setting and time variables of EPRB experiments cannot be defined on one probability space [15].

Thus, the postulate of counterfactual definiteness in conjunction with the use of a setting variable restricts the independent variables additional to $\boldsymbol{j}$ in the argument of Bell's functions $A$ to a, physically speaking, narrow subset of variables that we denote by $N_B$. This subset permits the physical description of static properties but cannot handle dynamic properties expressed by space-time dependencies.

As a consequence, the choices that can be made for variables in addition to the setting variable $\boldsymbol{j}$ in Bell's theory are extremely limited, particularly if these variables are related to space-time (or space and time). This limitation is so severe that it is impossible to describe general dynamic processes of classical physics with Bell's independent variables. The way to describe general dynamic processes in Kolmogorov's framework is by using stochastic processes.

To describe a dynamics of EPRB experiments one needs to use two dimensional vector stochastic processes, which involves several subtleties that, if neglected, lead to incorrect conclusions. A general vector stochastic process is in essence a vector of random variables, such a $(A_1(t_n), A_2(t_n), A_3(t_n), \cdots)$, whose statistical properties change in time (we use here discrete time only). A precise mathematical definition can be found in Ref. [16], pp. 11-15. In relation to EPRB experiments we thus consider vectors such as $(A_1(t_n), A_2(t_n))$.

A first difficulty that is usually encountered is related to the physics of spin measurements. According to Bohr, the outcomes of measurements on each separate side of the EPRB experiment are spin-up or spin down with equal likelihood, which appears to suggest stationarity or time-independence of the random variables $A_1(t_n)$ and $A_2(t_n)$. Bohr's postulate, however, does not necessitate a time-independence of the statistical correlations between the random variables. This fact has been explained on the basis of a mathematical model involving time in Ref. [5] (pp. 55-60) and demonstrated by actual EPRB related computer experiments [7].

A second difficulty arises from the fact, explained in detail above, that the time and setting related variables of EPRB experiments cannot be treated as independent. This difficulty can be resolved by use of the following two-dimensional system of functions (vector stochastic process) on a probability space $\Omega$:

$$\left(A_{\boldsymbol{j}_n}^{t_n}(\omega), A_{\boldsymbol{j}'_n}^{t'_n}(\omega)\right). \tag{2}$$

Settings and times are now included as indexes that are not independent. $\boldsymbol{j}_n = \boldsymbol{a}, \boldsymbol{b}$ represents the randomly chosen settings at one measurement place and $\boldsymbol{j}'_n = \boldsymbol{b}, \boldsymbol{c}$ at the second. $t_n$ as well as $t'_n$ are the respective measurement times. $n = 1, 2, 3, \cdots$ indicates just the number of the experiment. Only one setting can occur at one given time in order to avoid physical contradictions and incorrect assignments of probability measures. (Note that a generalization of the time-indexes to space-time $\boldsymbol{st}_n$ is straightforward.)

Bell's inequality then transforms to:

$$A_{\boldsymbol{a}}^{t_n}(\omega)A_{\boldsymbol{b}}^{t'_n}(\omega)+A_{\boldsymbol{a}}^{t_k}(\omega)A_{\boldsymbol{c}}^{t'_k}(\omega)-A_{\boldsymbol{b}}^{t_m}(\omega)A_{\boldsymbol{c}}^{t'_m}(\omega)\leq 3, \tag{3}$$

where the labels $n$, $k$, $m$ are the appropriate, all different, experiment numbers for which the particular settings have been chosen. Equation (3) puts no restrictions on the correlations of EPRB experiments, because the actual experiments may now be represented by a countable infinite number of different functions instead of the three or four functions used by Bell.

There do exist theorems that appear to prove the validity of Bell's inequality for stochastic processes (the Martingales discussed in [17] are just special forms of stochastic processes). These theorems, however, do not use two-dimensional vector stochastic processes as used in Equation (2). They use, instead, counterfactual definiteness in conjunction with setting variables to arrive at three-, four- or higher dimensional stochastic processes (Martingales). Thus these theorems cannot encompass dynamic measurement processes [18] and time- (space-time-) related variables, because they would then imply the existence of events with more than one setting at a given measurement time and, therefore, involve impossible events with non-zero probability measure. Such theorems apply, therefore, only to the set of variables $N_B$ as defined above and do not apply to EPRB types of experiments that may involve dynamical processes in the measurement equipment.

It is, therefore, imperative to view EPRB experiments in a different light. A violation of Bell-type inequalities need not be seen as crossing the border between the reasoning of classical Einstein type of physics and quantum mechanics, but indicating a possible dynamics in the interactions of particles and measurement equipment. This possible dynamics is what needs to be investigated, particularly as contrasted to the characterization of the measurement equipment by a completely static symbol [19].

## 5. Einstein Locality and Bell's Reasoning Revisited

Up to now, experimentalists have not used Bell's theorem and its implications to search for a many body dynamics of local equipment. Instead, they have attempted to "uncover" the instantaneous dynamic influences of remote measurements, the so called quantum non-localities. Some consider these non-localities to be the most profound development of modern physics [5]. They maintain that the measurement of the entangled partner causes instantaneous influences over arbitrary distances.

This search for influences due to distant events is based on the conviction, dating back to Bell's original paper, that Einstein locality is necessary to derive Bell's inequality. However, this is not the case. Bell's assumption that $\lambda$ is independent of the setting variable $\boldsymbol{j}$ is already contained in the postulate of counterfactual definiteness. The postulate of Einstein locality is not only redundant because of this fact, but does not require at all that $\lambda$ be independent of all settings. Variables dependent on the local setting and describing local many body interactions with the incoming particles are entirely permitted and necessary. It is counterfactual definiteness that requires that all additional variables represented by $\lambda$ be independent of the setting variable. But why does our classical theory need to involve the setting variable in the way Bell has included it? One can use the setting variable as an index together with another index related to or representing space-time. These indexes are, of course not independent as was pointed out above for stochastic processes.

From these facts, we can deduce that Einstein locality is not a necessary condition for Bell's derivation, rather the opposite. Its correct implementation prevents the derivation of Bell to go forward, as shown in Equation (3).

## 6. Conclusion

The major premise for the derivation of Bell's inequality is counterfactual definiteness, which in connection with Bell's use of setting variables restricts the domain of the variables in the argument of Bell's functions $A$ to a subset $N_B$ of general physical independent variables. $N_B$ does not encompass the variables that are necessary to describe a general dynamics of many body interactions with the measurement equipment. Using only the independent variables defined by $N_B$, it is impossible to find a violation of Bell's inequality, which therefore represents a demarcation between possible and impossible experience [9], not between classical and quantum physics. For a wider parameter space that permits the description of dynamic processes and includes space-time coordinates, the validity of Bell-type inequalities cannot be and has not been derived. This situation is reminiscent of that with the last theorem of Fermat before 1984. There existed only rather trivial proofs of Fermat's theorem for subsets of conditions, while a general proof was not known until Andrew Wiles supplied it in 1984. Such more complicated and general proofs of Bell's theorem have not been presented and, in the authors opinion, are not likely to be presented in the future, because they would need to remove the use of the setting variable $\boldsymbol{j}$.

## References

[1] Bell, J.S. (2001) On the Foundations of Quantum Mechanics. World Scientific, Singapore City, 228-222.

[2] Peres, A. (1995) Quantum Theory: Concepts and Methods. Kluwer Academic Publishers, Dordrecht, 2006-2007.

[3] Leggett, A.J. (1987) The Problems of Physics. Oxford University Press, Oxford, 164-165.

[4] Gilder, L. (2008) The Age of Entanglement. Alfred A. Knopf, New York, 269-296.

[5] Hess, K. (2015) Einstein Was Right! Pan Stanford Publishing, Singapore City.

[6] De Raedt, H., Jin, F. and Michielsen, K. (2013) *Proceedings of SPIE*, **8832**, N1-11. http://dx.doi.org/10.1117/12.2021860

[7] De Raedt, H., Michielsen, K. and Hess, K. (2016) Irrelevance of Bell's Theorem for Experiments Involving Correlations in Space and Time: A Specific Loophole-Free Computer-Example. arXiv:1605.05237 [quant-ph]

[8] Brans, C. (1987) *International Journal of Theoretical Physics*, **27**, 219-226. http://dx.doi.org/10.1007/BF00670750

[9] Boole, G. (1862) *Philosophical Transactions of the Royal Society of London*, **152**, 225-252. http://dx.doi.org/10.1098/rstl.1862.0015

[10] Vorob'ev, N.N. (1962) *Theory of Probability and Its Applications*, **7**, 147-162. http://dx.doi.org/10.1137/1107014

[11] Hess, K., De Raedt, H. and Michielsen, K. (2016) From Boole to Leggett-Garg: Epistemology of Bell-Type Inequalities. arXiv:1605.04887v1 [quant-ph]

[12] Aspect, A., Dalibard, J. and Roger, G. (1982) *Physical Review Letters*, **49**, 1804-1807. http://dx.doi.org/10.1103/PhysRevLett.49.1804

[13] Weihs, G., Jennewein, T., Simon, C., Weinfurther, H. and Zeilinger, A. (1998) *Physical Review Letters*, **81**, 5039-5043. http://dx.doi.org/10.1103/PhysRevLett.81.5039

[14] Khrennikov, A.Y. (2009) Contextual Approach to Quantum Formalism. Springer, Berlin. http://dx.doi.org/10.1007/978-1-4020-9593-1

[15] Hess, K. and Philipp, W. (2005) *Foundations of Physics*, **35**, 1749-1767. http://dx.doi.org/10.1007/s10701-005-6520-y

[16] Breuer, H.-P. and Petruccione, F. (2002) Open Quantum Systems. Oxford University Press, Oxford.

[17] Gill, Richard D. (2003) Accardi Contra Bell (Cum Mundi): The Impossible Coupling. Mathematical Statistics and Applications: Festschrift for Constance van Eeden, Institute of Mathematical Statistics, Beachwood, 133-154. http://projecteuclid.org/euclid.lnms/1215091935

[18] Allahverdyan, A.E., Balian, R. and Nieuwenhuizen, T.M. (2013) *Physics Reports*, **525**, 1-166. http://dx.doi.org/10.1016/j.physrep.2012.11.001

[19] Nieuwenhuizen, T.M. (2011) *Foundations of Physics*, **41**, 580-591. http://dx.doi.org/10.1007/s10701-010-9461-z

Journal of Modern Physics, 2016, 7, 1661-1667
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Influence of Recombination Centers on the Phase Portraits in Nanosized Semiconductor Films

**Gafur Gulyamov[1], Abdurasul G. Gulyamov[2], Feruza R. Muhitdinova[3]**

[1]Namangan Engineering-Pedagogical Institute, Namangan, Uzbekistan
[2]Physico-technical Institute, Academy of Sciences of Uzbekistan, Tashkent, Uzbekistan
[3]Namangan State University, Namangan, Uzbekistan
Email: gulyamov1949@mail.ru

**How to cite this paper:** Gulyamov, G., Gulyamov, A.G. and Muhitdinova, F.R. (2016) Influence of Recombination Centers on the Phase Portraits in Nanosized Semiconductor Films. *Journal of Modern Physics*, **7**, 1661-1667.
http://dx.doi.org/10.4236/jmp.2016.713151

**Received:** July 12, 2016
**Accepted:** September 11, 2016
**Published:** September 14, 2016



## Abstract

Influence of recombination centers' changes on the form of phase portraits has been studied. It has been shown that the shape of the phase portraits depends on the concentration of semiconductor materials' recombination centers.

## Keywords

Recombination Centers, Phase Portrait, Generation of Charge Carriers, Recombination of Charge Carriers, Forbidden Zone, Absorption Coefficient

## 1. Introduction

Operation recombination processes allow to control the number of excess charge carriers in semiconductor. Control of the concentration of charge carriers has special importance at the production of semiconductor devices. There are many types of recombination such as linear recombination and quadratic one, recombination through recombination centers and radiative recombination [1]-[3]. When recombination goes through recombination centers, the transition of the charge carriers from a free state to a bound one is independent of the presence of excess charge carriers of opposite sign. This means that there is no direct connection of an electron and the hole, that is, to complete the act first capturing one sign carrier by trap takes place and then capturing of the opposite sign does. There are many internal and external factors contributing to the growth and reduction of recombination centers. Changing the number of recombination centers may be undesirable for semiconductor devices and can disable them. In order to prevent interruption in the semiconductor devices work, it is necessary to conduct diagnostic of generation-recombination processes, and if it is necessary, replace them. The process of generation and recombination is similar to the oscillatory process, since the increase and the decrease of concentration of charge carriers occur periodically. For analysis of oscillatory processes, phase portraits (PhP) are used effectively, because they give the most complete picture of what is happening. In [4] using PhP the influence of frequency of variable deformation on the concentration of charge carriers in semiconductor illuminated by its forbidden zone's light, it is shown that the frequency of the variable deformation has a strong influence on the shape of the PhP. However, in [4] influence of recombination centers in the form of PhP is not investigated. In this paper, we investigate influence of the change of recombination centers in the semiconductor by PhP.

## 2. The Continuity Equation Taking into Account the Combined Effects of Light and Variable All-Round Deformation

The continuity equation expressing the change of concentration of charge carriers is described by the following expression

$$\frac{\partial n_e}{\partial t} = g - r + \frac{1}{q}\nabla \boldsymbol{I}_n \tag{1}$$

where $g$—rate of generation of charge carriers, $r$—rate of recombination of charge carriers, $q$—charge of electron, $\boldsymbol{I}_n$—the current density of electrons. In uniform sample, the continuity equation has the following form

$$\frac{\partial n_e}{\partial t} = g - r \quad (2).$$

when illuminated by its forbidden zone's light semiconductor becomes sensitive to external influences change of the absorption coefficient contributes to that [5]. At radiation generation is expressed by the following expression $g_s = \frac{\alpha I}{h\nu}$. Here, $\alpha$-absorption coefficient, $I$—intensity of the light, $h$—Planck's constant, $\nu$—the frequency of the light. With the express permission of the transition frequency dependence of the absorption coefficient is $\alpha = A\left(h\nu - E_g\right)^{1/2} = Ah^{1/2}\left(\nu - \nu_m\right)^{1/2}$ where $\nu_m$—the frequency of its forbidden zone's light [1] [3] [6] $A$-certain coefficient which is defined by the following expression $A = \frac{q^2\left(2m_{np}^*\right)^{3/2}}{nch^2 m_e^*}$ for direct allowed transitions $A \approx 2\times10^4$ [6], here $m_{np}^* = \frac{m_h^* m_e^*}{m_h^* + m_e^*}$—is the reduced effective mass, $m_e^*, m_h^*$—the effective mass of electrons and holes, respectively $n$—the index of refraction of the light, $c$—velocity of the light.

Let's consider the case when deformation of all-round strain effects on the semiconductor. When the deformation of band gap changes as follows [7]-[9] $E_g = E_g(0) - \Xi\varepsilon$, here $\Xi$—the constant of the deformation potential, $\varepsilon$—relative deformation. In the case $E_g(0) = h\nu$, then if $\nu = \nu_m$, $\left(h\nu - E_g(0) + \Xi\varepsilon\right)^{1/2} = \left(\Xi\varepsilon\right)^{1/2}$ the absorption coefficient becomes $\alpha = A\left(\Xi\varepsilon\right)^{1/2}$. If the deformation changes periodically $\varepsilon = \varepsilon_0\left(1 + \sin\omega_d t\right)$, whereas $\alpha = A\left(\Xi\varepsilon_0\left(1 + \sin\omega_d t\right)\right)^{1/2}$ here $\omega_d$ frequency of variable deformation, here generation at light $g_s = \frac{A\left(\Xi\varepsilon_0\left(1+\sin\omega_d t\right)\right)^{1/2} I}{h\nu}$. Equation (2) for this case is as follows:

$$\frac{\partial n}{\partial t} = g_0\left(1+\sin\omega_d t\right)^{1/2} - r\,,\quad g_0 = \frac{A\left(\Xi\varepsilon_0\right)^{1/2} I}{h\nu} \tag{3}$$

At direct unpermitted transition frequency dependence of the absorption coefficient $\alpha = B\left(h\nu - E_g\right)^{3/2} = Bh^{3/2}\left(h - h_m\right)^{3/2}$ where $\nu_m$—frequency of its forbidden zone's light, $B$—coefficient that is defined by the following expression $B = \frac{4}{3}\frac{q^2\left(2m_{np}^*\right)^{5/2}}{nch^2 m_e^* m_h^* h\nu}$, $B \approx 1.3\times10^4$ [6].

Equation (2) for direct unpermitted transition will be as follows:

$$\frac{\partial n}{\partial t} = g_0\left(1+\sin\omega_d t\right)^{3/2} - r\,,\quad g_0 = \frac{B\left(\Xi\varepsilon_0\right)^{3/2} I}{h\nu} \tag{4}$$

Taking into account the permitted and unpermitted direct transition continuity equation takes the form

$$\frac{\partial n}{\partial t} = \frac{A\left(\Xi\varepsilon_0\right)^{1/2} I}{h\nu}\left(1+\sin\omega_d t\right)^{1/2} + \frac{B\left(\Xi\varepsilon_0\right)^{3/2} I}{h\nu}\left(1+\sin\omega_d t\right)^{3/2} - r \tag{5}$$

Let's consider the case where the recombination through recombination centers takes place. According to statistics of the Shockley-Read, the rate of recombination is described by the expression: $r = N_t \frac{c_n c_p\left(pn - n_i^2\right)}{c_n\left(n+n_1\right)+c_p\left(p+p_1\right)}$ Here, $N_t$—the concentration of recombination centers, $c_n$ and $c_p$ coefficients capture electrons and holes, respectively [1] [3]. In this case, the continuity equation is

$$\begin{aligned}\frac{\partial n}{\partial t} = {} & \frac{A\left(\Xi\varepsilon_0\right)^{1/2} I}{\hbar\omega}\left(1+\sin\omega_d t\right)^{1/2} + \frac{B\left(\Xi\varepsilon_0\right)^{3/2} I}{\hbar\omega}\left(1+\sin\omega_d t\right)^{3/2} \\ & - N_t \frac{c_n c_p\left(pn - n_i^2\right)}{c_n\left(n+n_1\right)+c_p\left(p+p_1\right)}\end{aligned} \tag{6}$$

## 3. Analysis of Phase Portraits

Let's consider the effect of deformation on a illuminated semiconductor. Let's assume the following values: the temperature $T$ = 300 K, the band gap $Eg$ = 1.1 eV, the relative deformation $\varepsilon = 10^{-6}$, the deformation potential's constant is $\Xi$ = 11.4 eV, lightl's intensity $I = 10^{18}$ cm$^{-2}$·sek$^{-1}$, the effective mass off electron $m_n^*$ = 0.4·$m_0$, were $m_0$—mass of free electron, the effective mass of holes $m_p^*$ = 0.541·$m_0$, the coefficient of direct allowed transitions $A = 2\times10^4$ [6], the coefficient of unpermitted direct transition $B$ = $1.3\times10^4$ [6], the concentration of recombination centers $N_t = 2\times10^{13}$ cm$^{-3}$, the electron capture coefficient $c_n = 4.4\times10^{-10}$ cm$^3$/sek, hole capture coefficient $c_p = 6.2\times10^{-9}$ cm$^3$/sek, the intrinsic concentration $n_i = 2.2\times10^{13}$ cm$^{-3}$ [6] [10] [11]. Let the frequency of the variable all-round compression is $\omega_d$ = 250 Hz. Let's consider one period of variable strain. Let's consider the phase portrait in 0.3 seconds after the start of a periodic deformation. The beginning of counting out is $t = 0.3$ sek, because by this time the carrier concentration becomes the stable value, $n \approx 6\times10^{15}\ \text{cm}^{-3}$ in which will returns periodically at the process of variable deformation of the sample (see Figure 1 and Figure 2, point 1). As the deformation increases of the carrier concentration will increase and reach its peak $n \approx 1\times10^{16}\ \text{cm}^{-3}$ at point 2 as shown in Figure 1 and Figure 2. After this, decrease of the strain begins and the concentration at the point 3 has again significance $n \approx 6\times10^{15}\ \text{cm}^{-3}$. Further, when the strain is completely takes off the concentration takes minimum value $n \approx 10^{15}\ \text{cm}^{-3}$ at point 4. Upon completion of the period of hydrostatic compression at point 5 the concentration again returns to its primary value $n \approx 6\times10^{15}\ \text{cm}^{-3}$. This process will continue in this way until all parameters of the oscillatory system remain constant, and the phase portrait will take the form of a closed loop. $n \approx 10^{15}\ \text{cm}^{-3}$

At prolonged effect of the strain, especially if the deformation is variable there is the probability of increase of recombination centers. The appearance of new defects and structural changes in the semiconductor caused by fatigue and wear material promote it. Let's consider the effect of changes in the concentration of recombination centers in the form of phase portraits. Let the concentration of recombination centers grows increases from $2 \times 10^{13}$ cm$^{-3}$ to $8 \times 10^{13}$ cm$^{-3}$, in this case the phase portrait will not be in the form of a closed loop, but it will curl into a spiral form (see Figure 3). The maximum value of the carrier concentration will be $n \approx 1.6\times10^{16}\ \text{cm}^{-3}$ when $N_t = 2\times10^{13}\ \text{cm}^{-3}$, and at $N_t = 8\times10^{13}\ \text{cm}^{-3}$ the maximum concentration of charge carriers will be

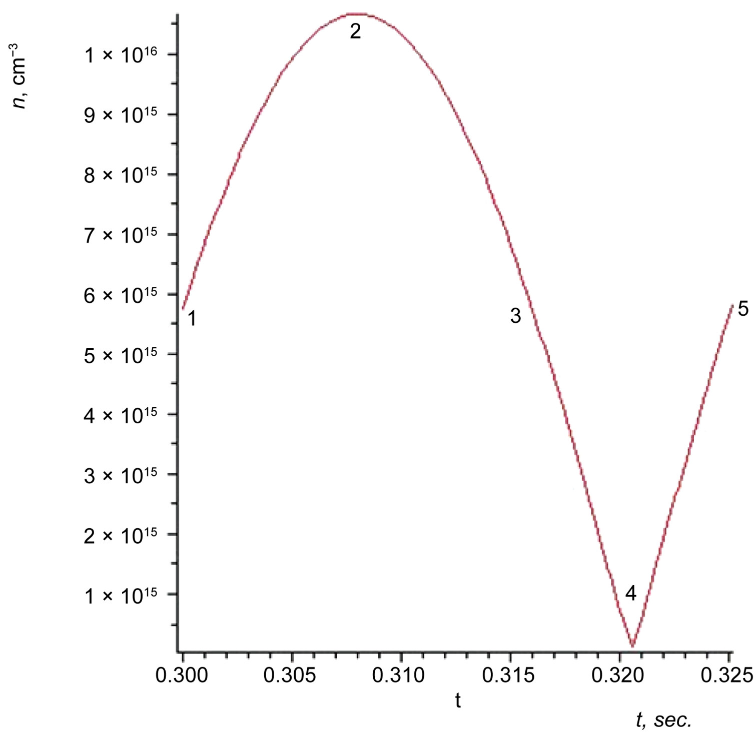



Figure 1. Dynamics of changes of the concentration of charge the carriers in the period of variable strain.

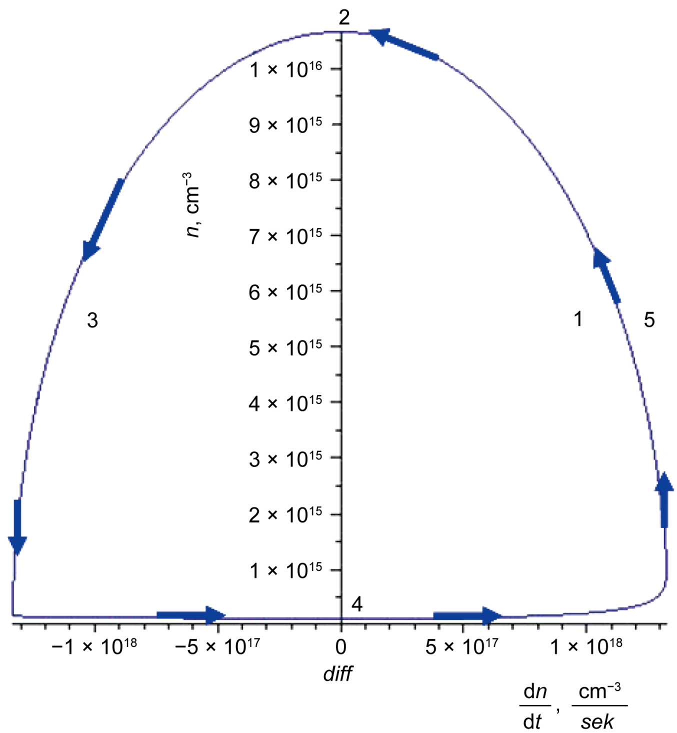



**Figure 2.** Phase portrait of the concentration ($n$) of charge carriers versus the rate of change of the charge carrier concentration ($\mathrm{d}n/\mathrm{d}t$).

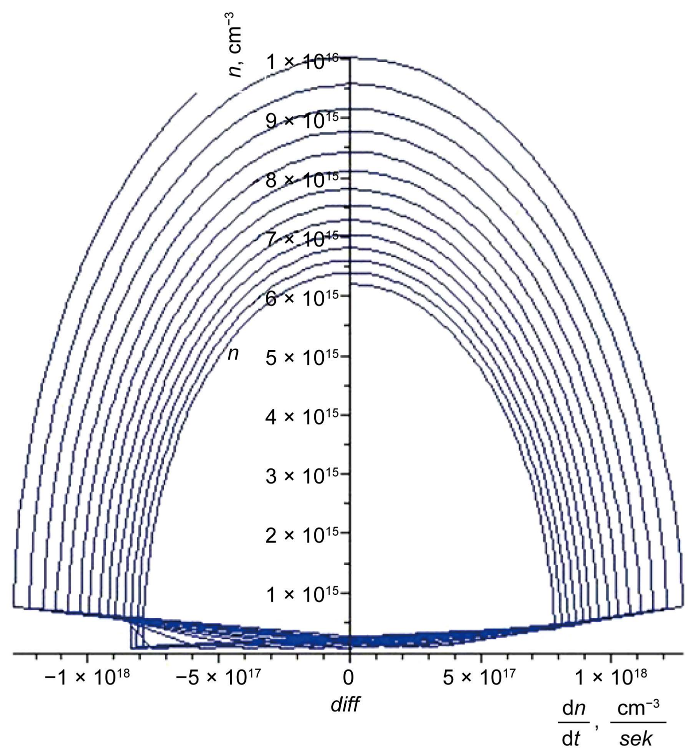



**Figure 3.** Phase portrait $n$ versus $\mathrm{d}n/\mathrm{d}t$, for the case when the concentration of recombination centers variers from $N_t = 2\times10^{13}\,\mathrm{cm}^{-3}$ to $N_t = 8\times10^{13}\,\mathrm{cm}^{-3}$.

$n \approx 6\times10^{15}\ \text{cm}^{-3}$. Such decrease of the concentration of charge carriers can be explained by the fact that the generation changes in a constant range, and the range of the recombination increases with time.

Figure 4 shows the phase portrait for the case when the concentration of recombination centers decreases from 2 × $10^{13}$ $\text{cm}^{-3}$ to 2 × $10^{12}$ $\text{cm}^{-3}$. In this case, when $N_t = 2\times10^{13}\ \text{cm}^{-3}$ the maximum concentration is $n \approx 1.6\times10^{16}\ \text{cm}^{-3}$, and when $N_t = 2\times10^{12}\ \text{cm}^{-3}$ the maximum concentration is $n \approx 2.9\times10^{16}\ \text{cm}^{-3}$. Increasing concentration of charge carriers is caused by that the range of the recombination term $R$ in the Equation (2) decreases but the range of generation term $g$ remains constant.

Consideration of the phase picture's transformation at the change of the system parameters is very important for understanding the physical processes in the system. Looking at the "phase portrait" under certain given values of the parameters it is possible to imagine all the possible movements in the system for any initial values. While you observe the modification in the picture at the change of the parameters, you represent all advances that the given physical system can have for all possible values of the parameters. For example, the location and nature of the singular points on the phase plane make possible to do a number of conclusions about the processes in the system.

Research of generational process by phase portraits method has practical importance because the phase portraits give the most complete picture of what is happening in the semiconductor. This allows make diagnostics of semiconductor devices and to replace them timely for preventing interruption of their work.

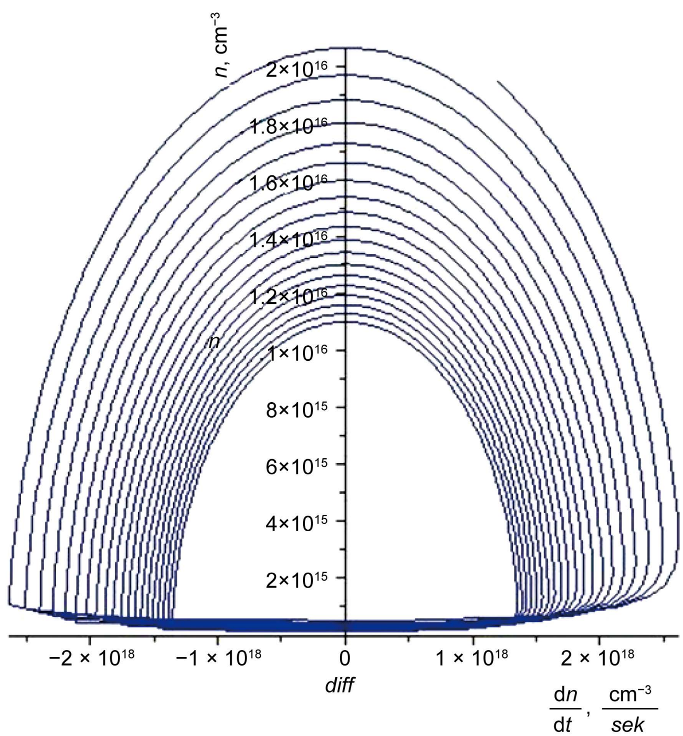



Figure 4. Phase portrait $n$ versus d$n$/d$t$, for the case when the concentration of recombination centers variers from $N_t = 2\times10^{13}\ \text{cm}^{-3}$ to $N_t = 2\times10^{12}\ \text{cm}^{-3}$.

## 4. Conclusions

Thus, the study of generation-recombination processes on the basis of the phase portraits allows us to make the following conclusions:

- Increasing recombination centers in the semiconductor causes decrease of carrier concentration, and the phase trajectory rolls spirally toward the lowest values, both on the axis $n$ and on the axis $\mathrm{d}n/\mathrm{d}t$.
- Decreasing of recombination centers in the semiconductor causes with increasing carrier concentration, and the phase trajectory is set in a spiral towards the largest values as along the axis $n$ as along the axis $\mathrm{d}n/\mathrm{d}t$.
- The phase portraits permitted to conduct diagnostics of semiconductor devices

## Acknowledgements

This work was supported by government grants of Uzbekistan F2-OT-O-15494 "Improving the efficiency of the radiators, solar cells and other optoelectronic devices based on quantum dot research and hetero nanostructure exciton and bipolaron transport phenomena", F2-FA-0-97004 "Dynamics of solitons in ultracold quantum gases, and optical systems and synergistic processes in multicomponent semiconductors" and F2-21 "Mathematical modelling of the determination of the density of surface states at the semiconductor-insulator".

## References

[1] Bonch-Bruevich, V.L. and Kalashnikov, S.G. (1977) Semiconductor Physics. Nauka, Moscow, 674 p.

[2] Abakumov, V.N., Perel, V.I. and Yassievich, I.N. (1991) Nonradiative Ecombination in Semiconductors. In: Agranovich, V.M. and Maradudin, A.A., Eds., *Modern Problems in Condensed Matter Sciences*, North-Holland, Amsterdam, 376 p.

[3] Kireev, P.S. (1975) Semiconductor Physics. Vishaya Shkola, Moscow, 584 p.

[4] Gulyamov, G., Gulyamov, A.G. Ergashev, A.Q. and Abdulazizov, B.T. (2015) *Journal of Modern Physics*, **6**, 1921-1926. http://dx.doi.org/10.4236/jmp.2015.613197

[5] Gulyamov, G. and Gulyamov, A.G. (2015) *Semiconductors,* **49**, 819-822. http://dx.doi.org/10.1134/S1063782615060111

[6] Pankove, J.I. (1971) Optical Processes in Semiconductors. Prentice-Hall, Inc., Englewood Cliffs, 457 p.

[7] Polyakova, A.L. (1979) Deformation of Semiconductors and Semiconductor Devices. Energy, Moscow, 168 p.

[8] Gulyamov, A.G. (2015) *World Journal of Condensed Matter Physics*, **5**, 60-65. http://dx.doi.org/10.4236/wjcmp.2015.52009

[9] Bir, G.L. and Pikus, G.E. (1972) Symmetry and Strain-Induced Effects in Semiconductors. Nauka, Moscow, 584 p.

[10] Baranskiy, P.I., Klotchkov, V.P. and Potykevich, I.V. (1975) Semiconductor Electronics. Naukova Dumka, Kiev.: 704 p.

[11] Bonch-Bruevich, V.L., Zvyagin, I.P., Karpenko, I.V. and Mironov, A.G. (1987) Problems in Physics of Semiconductors. Nauka, Moscow, 144 p.

Journal of Modern Physics, 2016, 7, 1668-1701
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Black Body Quantum Fluctuations and Relativity

**Sebastiano Tosto**

Retired Physicist, ENEA Casaccia, Roma, Italy
Email: stosto44@gmail.com, stosto@inwind.it

**How to cite this paper:** Tosto, S. (2016) Black Body Quantum Fluctuations and Relativity. *Journal of Modern Physics*, **7**, 1668-1701.
http://dx.doi.org/10.4236/jmp.2016.713152

**Received:** July 20, 2016
**Accepted:** September 16, 2016
**Published:** September 19, 2016



## Abstract

The paper introduces a simple theoretical model aimed to provide a possible derivation of the quantum fluctuations of the black body radiation. The model offers the chance of inferring and linking contextually quantum and relativistic results.

## Keywords

Quantum Physics, Thermodynamics, Relativity

## 1. Introduction

In 1859, Kirchhoff had the remarkable idea that a small hole in the side of a massive body of material containing a large cavity was the best experimental approximation of the concept of total absorber: the radiation penetrating through the hole was correctly assumed bouncing between the internal walls of the cavity with a little probability of escaping outside. With this viewpoint, still today acknowledged [1], Planck modeled 1901 the thermodynamic equilibrium of the radiation field inside the cavity. Since any thermodynamic system is subjected to statistical fluctuations around the equilibrium configuration, Einstein proposed in 1909 a theoretical model about these fluctuations working on the Planck result. The Einstein model was focused essentially on the black body radiation assumed at the equilibrium in a cavity with perfectly reflecting walls. This assumption arose however the difficulty of explaining the thermalization mechanism of the radiation field. The second law of thermodynamics states that any system left undisturbed for a sufficiently long time tends to the equilibrium state [2]; nevertheless the thermalization time of photons at temperatures below 109 K is expectedly very long, as their direct interaction is negligible compared to that with matter [3]. The fact that the thermalization process is slightly shortened in the presence of rarefied gas particles [4], shows that in fact the interaction of photons with matter, *i.e.* with the internal walls of the cavity, is required to explain the equilibrium condition of the black body at the usual temperatures and times at which is tested the Planck law. The equilibrium condition is attained therefore considering partially reflecting walls of the cavity to promote the photon-solid matter interaction mechanism via continuous absorption and reemission of radiation.

The problem of the quantum fluctuations of black body radiation is still today debated for its theoretical interest [5] [6], in particular as concerns the thermalization mechanism of the photons in the cavity. Just this is the problem: any model aimed to describe the Planck law and its transient deviations from the equilibrium should infer explicitly this kind of interaction, without need of postulating it separately and purposely. Moreover, the fluctuation of a thermodynamic system implies in general several non-equilibrium phenomena, e.g. local temperature gradients and configuration changes; specifically, are expected gradients of radiation frequency and mass evaporated from the internal surface of the cavity, whose dynamics contributes to the thermalization of photons.

While focusing on the radiation field only seems reductive, the variety of phenomena involved when a black body system is out of the equilibrium suggests the usefulness of a comprehensive approach to the problem and introduces the three main motivations of this paper:

1) To propose a model where the photon interaction with the walls of the cavity appears as a natural consequence of the theoretical approach underlying the black body physics.

2) To highlight the thermodynamic aspects of the black body fluctuations with reference to their quantum basis, in particular the uncertainty principle.

3) To show that relativistic results are also obtainable in the frame of a unique conceptual model.

After a preliminary outline of the main dynamical variables prospectively implicated in the problem, the model is specifically addressed to introduce not only the fluctuation but also the main physical laws expectedly useful to describe it. Despite the inherent complexity of the problem, the exposition is organized in order to be as simple, gradual and self-contained as possible.

## 2. Preliminary Considerations

Consider one free particle of mass $m$ moving within a space range $\Delta x$ during a time range $\Delta t$. It is in principle possible to express $\Delta x$ as a function of the Compton length $\lambda_C$ of the particle; so define the range size in $\lambda_C$ units putting

$$\Delta x = n\frac{h}{mc}, \quad n \geq 1, \tag{1}$$

where $n$ is an arbitrary real number. The second position emphasizes that the range $\Delta x$ where the particle is allowed to move cannot be smaller than $\lambda_C$, which is an intrinsic physical property of the particle itself through its mass $m$.

In principle, nothing hinders to express the numerical parameter $n$ as the ratio $c/v$, being $v$ the component of velocity of the particle along $\Delta x$: is attracting the chance that $n \to 1$ is consistent with $\Delta x \to \lambda_C$ whereas $\Delta x \to \infty$ for $v \to 0$, in agreement with the arbitrary size of $\Delta x$ and value of $v$ in a given reference system *R*. The formal choice of introducing the range size in $\lambda_C$ units does not exclude thinking even a photon confined in $\Delta x$ during a time range $\Delta t$, which compels therefore defining $\Delta x = c\Delta t$. Dividing both sides of Equation (1) by $c\Delta t$, one finds in $R$

$$\frac{\Delta x}{c\Delta t} = n\frac{h}{mc^2\Delta t}.$$

Defining the link between time and space range sizes via *c*, in order to ensure that any massive particle is effectively confined within $\Delta x$ during $\Delta t$, the positions

$$\frac{\Delta x}{\Delta t} = c, \quad \nu = \frac{1}{\Delta t}, \quad v \le c \tag{2}$$

yield

$$\varepsilon_0 = nE, \quad n = \frac{c}{v}, \quad \varepsilon_0 = mc^2, \quad E = h\nu. \tag{3}$$

The first equation simply rewrites $\varepsilon_0 \ge E$ as $mc^2 = h\nu_m$, being $\nu_m = c/(v\Delta t) \ge \nu$.

Simple considerations show that these positions are physically sensible. The first equation reads indeed

$$mv = \frac{h\nu}{c} = \frac{h}{\lambda}, \quad h\nu = mc^2\frac{v}{c}, \quad \lambda = \frac{c}{\nu}, \tag{4}$$

*i.e.* $mv$ is related to the wavelength $\lambda$. Regard $\lambda$ with the physical meaning of wavelength of matter wave introducing a multiplicative factor $\gamma < 1$ necessary to express the wavelength as a function of v instead of *c*, as it appears in the third equation. Divide both sides of the first equation by the arbitrary number $\gamma$; one obtains

$$\frac{mv}{\gamma} = \frac{h\nu}{\gamma c} = \frac{h}{\lambda'}, \quad \lambda' = \frac{\gamma c}{\nu} = \gamma\lambda, \quad \gamma < 1, \tag{5}$$

so that $\lambda'$ is in effect defined by $v' = \gamma c$ instead of *c*; then

$$\frac{h}{\lambda'} = p = \frac{mc^2}{\gamma}\frac{v}{c^2}$$

yields

$$\frac{h}{\lambda'} = p = \varepsilon\frac{v}{c^2}, \quad \varepsilon = \frac{mc^2}{\gamma} = \frac{\varepsilon_0}{\gamma}, \quad p = \frac{mv}{\gamma}. \tag{6}$$

In other words, the matter wave propagating at rate $v < c$ implies $\lambda' < \lambda$ in agreement with $\varepsilon_0 > E$. The result (6) is interesting because it is easy to show that

$$\gamma = \sqrt{1-(v/c)^2}; \tag{7}$$

so, through the position $n = c/v$, Equation (6) yield the De Broglie wave momentum of the particle and contextually the relativistic expressions of energy and momentum, whereas Equations (1) and (6) imply the corpuscle/wave dual behavior of matter.

To check this point, replace in the first Equation (6) $\nu' = v/\lambda'$, so that $v^2/c^2 = h\nu'/\varepsilon$

and thus $\varepsilon\left(1-v^2/c^2\right)=\varepsilon-h\nu'$ ; hence

$$-mc^2\sqrt{1-v^2/c^2}=L=h\nu'-\varepsilon.$$

At the left hand side appears the Lagrangian $L$ of the free particle, which correctly results as a difference of two energies. Indeed by definition $h\nu'=(h/\lambda')v=\boldsymbol{pv}$ , so that $\varepsilon=\boldsymbol{pv}-L$ . The concept of action $\int Ldt$ also follows by consequence. Note that in the expressions (6) of $\varepsilon$ and $p$ appears the ratio $m/\gamma$ ; admitting that this ratio is finite even in the limit $v\rightarrow c$ if contextually $m\rightarrow 0$, thus obtaining the undetermined form 0/0, one finds $\varepsilon=m_0c^2$ and $p=m_0c$ having put by definition

$$m_0=\lim_{\substack{v\to c\\ m\to 0}}\left(m/\gamma\right). \tag{8}$$

This limit holds for a photon wave, in which case Equation (6) yields $\varepsilon=cp$ . For a massive particle instead

$$p=\frac{mv}{\gamma}=\frac{h\nu}{c\gamma}=\frac{h}{c}\frac{1}{\Delta t'},\quad \Delta t'=\Delta t\sqrt{1-v^2/c^2}. \tag{9}$$

If $mv/\gamma$ is an invariant expression of momentum, then the right hand side must be an invariant quantity as well; in effect $\Delta t'$ for $v=const$ is the Lorentz transformation of $\Delta t$ between inertial reference systems displacing at rate v, in either of which the particle is at rest. Moreover it appears that $\gamma$ is not mere numerical factor, actually it allows linking the cases $m\neq 0$ and $m=0$ depending on whether $v/c\leq 1$.

These conclusions are inferred regarding in particular $\Delta x$ as a mere range size; *i.e.* the kinetic properties of a free particle follow simply as a consequence of the space and time ranges available to and compatible with its dynamical behavior. The frequency $\nu$ defines the range size

$$\Delta x=\frac{c}{\nu} \tag{10}$$

according to Equation (2); *i.e.*, in agreement with the dual behavior of matter, the range size is related via $\nu$ to the wavelength of the pertinent matter wave.

Furthermore, an interesting consequence follows regarding $\Delta x$ as a physical constrain to the particle delocalization: for example one could suppose that $\Delta x$ is delimited by two infinite potential walls that define its boundaries, in which case the particle must be thought bouncing back and forth in a given space range without chance of escaping. In other words, Equation (1) does not exclude that t the time $\Delta t+\delta t$ the particle could be located at $\Delta x+\delta x$ , as instead it is purposely excluded now. If so, then $\Delta x$ is actually an one-dimensional cavity; thus the concept of frequency $\nu$ is no longer the reciprocal time range $\Delta t^{-1}$ necessary for the photon to travel $\Delta x$ , rather it is related to the bouncing rate physically implied by the boundary potential walls. This is understandable thinking a steady photon wave with wavelength $\lambda_{\max}=2\Delta x$ or matter wave with wavelength $\lambda'_{\max}=2\gamma\Delta x$ of Equation (5), both additional to all wavelengths allowed in the cavity. So, owing to Equation (2), the lowest frequencies allowed for photon or massive particle traveling through $\Delta x$ are respectively with obvious notation

$$\nu_{ph}^{0} = \frac{c}{\lambda_{\max}} = \frac{1}{2}\frac{1}{\Delta t}, \quad \nu_{m}^{0} = \frac{v}{\lambda'_{\max}} = \frac{1}{2}\frac{v/c}{\gamma\Delta t} = \frac{v/c}{\gamma}\nu_{ph}^{0}; \tag{11}$$

Consequently the minimum energies $h\nu/2$ and $h(v/\gamma\Delta x)/2$ regard $\Delta x$ as a one-dimensional cavity where matter particles or even photons are confined without chance of escaping. Note that $\lambda_{\max}$ and $\lambda'_{\max}$ have been inferred in Equation (11) after having simulated a confinement mechanism constraining any particle to move within $\Delta x$; it is easy to show however the possibility of reversing this path, *i.e.* that once admitting the existence of the limit momentum wavelengths $\lambda_{\max}$ and $\lambda'_{\max}$ it is possible to infer a mechanism that constrains the motion of any particle within $\Delta x$ only. This point is highlighted just below and later in the Section 7.

Implement to this purpose the case of a particle bouncing elastically back and forth against either boundary wall that delimits the confinement range, Equation (11); the momentum change of the particle reads thus $\Delta p = 2mv/\gamma$. If the bouncing lasts a time range $\Delta t$, the force acting on the wall is

$$\frac{\Delta p}{\Delta t} = F = \frac{2mvc}{\gamma\Delta x_{pw}};$$

the subscript $pw$ stands for "potential wall" to stress that this particular range is able to confine any particle. It is clearly possible to express $F$ in Planck units via an appropriate multiplicative factor $q$; then the last result reads

$$F = q\frac{c^4}{G} = \frac{2mvc}{\gamma\Delta x_{pw}},$$

which yields

$$\Delta x_{pw} = \frac{2mvG}{q\gamma c^3}.$$

This is not a hypothesis "ad hoc", as the Planck units have fundamental worth, being based on dimensional relationships involving fundamental constants of nature. It is immediate to describe in this respect the particular case of photon confinement taking the limit for $v \to c$ and $m \to 0$, which yields $m/\gamma \to m_0$ according to Equation (8). Putting then $M = m_0/q$ this limit corresponds to the confinement of a photon in $\Delta x$ and reads

$$\Delta x_{co} = \frac{2MG}{c^2}, \tag{12}$$

which expresses the condition even for a photon to be trapped inside any $\Delta x$ of such size together with $M$ by consequence of the gravitational effect of this latter. For obvious reasons, the subscript $pw$ has been replaced by that stressing the idea of $M$ driven confinement.

Start eventually from the identity (1) $\Delta x \equiv h/mv$ to obtain $p\Delta x = h/\gamma$ thanks to Equations (1) and (3); being by definition $\gamma < 1$, one infers $p\Delta x > h$ whatever $v \neq 0$ might be in the reference system $R$. Moreover, replacing $p$ via the first Equation (6) as well, one also finds $p\Delta x = h/\gamma = \varepsilon v\Delta x/c^2$; thus $h/\gamma = \varepsilon v\Delta t/c$ yields $(c/v)h/\gamma = \varepsilon\Delta t$ in the same *R*. As of course $c/v\gamma > 1$, reasoning exactly as before one finds $\varepsilon\Delta t > h$.

Consider now the identity $p \equiv (p+p_o) - p_o$, being $p_o$ a constant momentum component, and note that the right hand side defines an arbitrary range $\Delta p$ whose upper and lower boundaries are $p+p_o$ and $p_o$ respectively. Is essential the fact that both *p* and $p_o$ are arbitrary and independent each other, so that the same holds for the range size and its boundary coordinates. Regarding in an analogous way even $\varepsilon\Delta t$, the straightforward conclusion is

$$\Delta p\Delta x > h, \quad \Delta\varepsilon\Delta t > h.$$

Apart from the simplicity of reasoning, is remarkable the fact that the most typical feature of the quantum physics, the Heisenberg inequalities, has been obtained from the relativistic Equation (6).

Equation (3) and other results of this section have been inferred directly from general considerations about the properties of the space time [7] in the frame of a unique and comprehensive approach "*ab initio*". Equation (7) will be examined further on in the Sections 4 and 7 to clarify how these considerations are linked to the quantum fluctuations.

## 3. Fluctuations

This section introduces the fluctuation of all variables previously introduced, with the aim of finding possible links between these variations. Differentiate $\varepsilon_0 = nE$ to simulate the physical idea that both energies are subjected to fluctuate: as by definition the dynamical variable of $\varepsilon_0$ is the mass *m* whereas that of *E* is the frequency $\nu$, write thus according to Equation (3)

$$\delta\varepsilon_0 = n\delta E - \frac{E}{c}n^2\delta v, \tag{13}$$

being

$$\delta\varepsilon_0 = c^2\delta m, \quad \delta E = h\delta\nu. \tag{14}$$

These positions relate in general $\delta m$ and $\delta\nu$ to $\delta v$. Thinking specifically the black body cavity, for example, the first Equation (14) describes the fluctuations of the amount of mass evaporated from the internal walls of the cavity, the second Equation (14) concerns the corresponding frequency fluctuation of the radiation field in it contained. Equation (13) relates them and requires temperature fluctuation too, although not yet explicitly concerned. To clarify this point note that the changes $\delta\varepsilon_0$ and $\delta E$ are defined around the respective $\varepsilon_0$ and *E*, which can be regarded as equilibrium values. For the following purposes it is useful to calculate the average fluctuations $\langle\delta\varepsilon_0\rangle$ and $\langle\delta E\rangle$ considering arbitrary fluctuation ranges around an arbitrary reference energy value. For instance $\langle\delta E\rangle$ is calculated considering various ranges $\delta E_j$ of values around the equilibrium value $E_{eq}$ and taking their mean value; if $1 \le j \le N$, then $\langle\delta E\rangle = N^{-1}\sum_j \delta E_j$ is the average fluctuation of the system matter+radiation; the same holds indeed to define $\langle\delta\varepsilon_0\rangle$. In principle therefore $\langle\delta\varepsilon_0\rangle$ and $\langle\delta E\rangle$ are independent of the respective $E_{eq}$ and $\varepsilon_{0eq}$, which will be denoted in the following as $E$ and $\varepsilon_0$ for simplicity of notation. Trivial manipulations of Equation (13) yield

$$\delta\varepsilon_0 = n^2 Y \delta E = \zeta \delta E, \quad \zeta = n^2 Y \tag{15}$$

where

$$Y = \frac{1}{4n^2}\frac{\delta E}{\delta\varepsilon} - \left(\sqrt{\frac{\delta\varepsilon}{\delta E}} - \frac{1}{2n}\sqrt{\frac{\delta E}{\delta\varepsilon}}\right)^2, \quad \delta\varepsilon = \frac{E\delta v}{c}. \tag{16}$$

In general

$$Y = Y(\nu, r, v, \delta v), \quad r = \frac{\delta m}{\delta v} \tag{17}$$

emphasizes that the variables of the problem are four and that only $Y$ depends upon $\delta v$. Regard first $Y$ and $n^2$ separately and introduce the mean value $\langle Y \rangle$ calculated in arbitrary ranges of all variables except $\delta v$. Thus

$$\langle Y \rangle = \frac{1}{\Delta v}\frac{1}{\Delta \nu}\frac{1}{\Delta m}\iiint Y \mathrm{d}v \mathrm{d}\nu \mathrm{d}m$$

defines

$$k = k(\delta v) = \frac{1}{\langle Y \rangle}; \tag{18}$$

of course $\Delta v = v_2 - v_1$ denotes an arbitrary range of velocity components with respect to which is integrated $Y\mathrm{d}v$, whereas $\delta v$ is the velocity fluctuation concurring together with $\delta m$ and $\delta \nu$ to define the relationship (13) between $\delta E$ and $\delta\varepsilon_0$. The notation emphasizes that $k$ is actually an arbitrary numerical parameter, *i.e.* a scale factor dependent on $\delta v$ only, such that for example

$$k\delta E = \delta(kE), \quad k\delta\varepsilon_0 = \delta(k\varepsilon_0);$$

by consequence $k$ is also a conversion factor such that $kE$ and $k\varepsilon_0$ can be related to energies with different physical meaning with respect to the initial $E$ and $\varepsilon_0$.

Among the possible values of $Y$, calculate Equation (15) with the specific value $\langle Y \rangle$; hence, owing to Equations (3) and (13),

$$\frac{\delta E}{\delta\varepsilon_0} = \frac{k}{n^2} = k\frac{E^2}{\varepsilon_0^2}, \quad Y \equiv \langle Y \rangle. \tag{19}$$

This result yields:

$$\left\langle \frac{\delta E}{\delta\varepsilon_0} \right\rangle = k\left\langle \frac{1}{n^2} \right\rangle = \frac{k}{c^2}\langle v^2 \rangle = k\left\langle \frac{E^2}{\varepsilon_0^2} \right\rangle \tag{20}$$

These equations are obtained simply averaging the ratios of Equation (19). It worth emphasizing that $\langle v^2 \rangle$ is linked to the average temperature in the case of an ideal gas where by definition the particles are non-interacting; this shows that $v$ is the velocity of matter particles evaporated from the walls of the cavity. So $\langle v^2 \rangle$ is related to the equilibrium temperature of the cavity containing the black body radiation field. In this one-dimensional approach $\Delta x$ is the distance between two matter boundary surfaces confining $m$ and the photon wave of frequency $\nu$, both subjected to the respective fluctuations $\delta m$ and $\delta\nu$. Also, $\delta\nu$ is the obvious consequence of the expected temperature fluctuation around the average value pertinent to $\langle \nu^2 \rangle$.

So far the first Equation (3) is the only equation correlating $E$ and $\varepsilon_0$. To find a second equation, the one linking $E$ and $\nu$, consider now Equation (13) that involves the variables appearing in (17) and yields

$$n-\frac{E}{c}n^2\frac{\delta\nu}{\delta E}=\zeta,\quad \zeta=\zeta(E),\quad \nu=\nu(E) \tag{21}$$

It is evident that a hypothesis has been introduced regarding $Yn^2$ of Equation (15) as the function $\zeta(E)$. In principle Equation (21) is not new with respect to Equation (15); however the form of this latter was useful to infer Equation (18) and the temperature implied by Equation (20), whereas Equation (21) is now implemented thanks to its analytical form easily integrable. The solution

$$\nu=\frac{E}{Z_E+p_o},\quad Z_E=\int(\zeta/c)\mathrm{d}E \tag{22}$$

is the sought second equation linking $E$ and $\nu$; the notation emphasizes that the integration constant $p_o$ has physical dimensions of momentum. Put

$$\zeta=\zeta_0+\zeta_1 E+\zeta_2 E^2+\cdots, \tag{23}$$

which yields at the first order of approximation

$$\nu=\frac{E}{p_o+\zeta_0 E/c+\zeta_1 E^2/2c},\quad p_o=const'; \tag{24}$$

if the position (23) is correct, then even this lowest order of approximation should give a sensible result. The validity of Equation (24) is preliminarily proven recalling Equations (3), according which $\nu$ yields

$$n=\frac{c}{\nu}=\frac{p_o c}{E}+\zeta_0+\frac{\zeta_1 E}{2}$$

and then

$$\varepsilon_0=nE=p_o c+\zeta_0 E+\frac{\zeta_1 E^2}{2};$$

so, neglecting preliminarily the third addend at the right hand side, this result reads $\varepsilon_0=E+E_0$, having put $\zeta_0=1$ and $cp_o=E_0$. With the integration constant $p_o\geq 0$, therefore, the result is nothing else but the statement $\varepsilon_0\geq E$ of Equation (1). Clearly with an appropriate choice of the integration constant this inequality holds even retaining the $\zeta_1$ term.

Before proceeding, it is useful to verify further the validity of the equations hitherto inferred, in particular as concerns the physical meaning of the series expansion (23) of Equation (15). A simple one-dimensional approach is still enough for the present purposes.

## 4. Check of the Preliminary Results

Recalling Equations (10) and (1), trivial manipulations show that Equation (19) reads

$$\delta E = h\delta\nu = D\nu\delta m, \quad D = k\lambda_C^2\nu. \tag{25}$$

The first equation links $\delta m$ and $\delta\nu$: if in particular $\delta E$ is due to the change of photon frequency in $\Delta x$, then it can be nothing else but the quantum fluctuation of the radiation field in the range; also, $\delta\nu$ is related to the mass fluctuation rate $\nu\delta m$ occurring at a typical length scale of the order of the Compton length $\lambda_C$ that defines $\delta E$ through *D*. The free parameter *k* fits the basic physical definition $\lambda_C^2\nu$ of diffusion coefficient to the appropriate value in specific situations. Specifically, as it will be shown below, this result also implies regarding $\Delta x$ as the distance separating the surfaces of two bodies of matter: thinking for example to the black body, *m* can be the mass of a particle evaporated from the internal surface of the cavity and diffusing throughout the cavity, whose size $\Delta x$ is defined as a function of *n*.

From Equation (25) follow interesting consequences. Rewrite

$$\delta E = \frac{D}{\beta}\delta\log\left(m/V_o\right), \quad \beta = \frac{1}{m\nu}, \tag{26}$$

where $V_o$ is an arbitrary constant volume.

Note that $(m\nu)^{-1}$ has physical dimensions time/mass; thus $\beta$ is the particle mobility, also defined as velocity/force. Moreover $D/\beta$ has physical dimensions force × length, *i.e.* pressure × volume.

The dimensional analysis suggests that $D/\beta$ should be related to, and thus proportional to, $k_BT$. Putting indeed $D/\beta \propto k_BT$ and merging the proportionality constant with k of Equation (18), one finds concurrently three relevant results.

First the well known law $\text{pressure} \times \text{volume} = k_BT$; of course this result holds for non interacting particles, as in the case of an ideal gas, whereas *T* is clearly linked to $\langle v^2 \rangle$ previously found. With specific reference to the present model, the gas is that formed by evaporation of matter from the internal walls of the cavity containing the Planck radiation; $\langle v^2 \rangle$ is related to the temperature of gas particles in equilibrium with the surface of the cavity.

Moreover

$$\frac{D}{\beta} = k_BT, \quad \delta E = \delta\mu, \quad \mu = k_BT\log\left(\frac{C}{C_o}\right), \quad C = \frac{m}{V_o} \tag{27}$$

where $C$ is the concentration of *m* in $V_o$ and $C_o = m_o/V_o$ is a constant. So the former equation is the well known Einstein equation linking diffusion coefficient and mobility.

Eventually, the third equation defines the chemical potential; this clarifies the physical meaning of $\delta E$ and suggests the chance of identifying $C_o$ as the equilibrium uniform concentration that implies $\mu = 0$ in correspondence to $\delta m = 0$, which indicates the end of the diffusion process.

In effect, the diffusion equations are contextually obtainable. Dividing both sides of Equation (25) by $\delta x$, one finds

$$\frac{\delta E}{\delta x}=\frac{\delta\mu}{\delta x}=Dv\frac{\delta m}{\delta x}$$

*i.e.* this equation defines a force $F$ acting on $m$. It is easy to convert force into mass flux $J$, having physical dimensions of mass per unit time and surface, dividing both sides by $V_o v$. Recalling that by definition $F=-\delta\mu/\delta x$, this equation reads

$$J=-D\frac{\delta C}{\delta x},\quad J=\frac{F}{V_o v}.$$

This is the well known Fick diffusion law, from which also follows the second Fick law with the help of an appropriate continuity equation that excludes mass sinks or sources within $V_o$. Given a function $f=f(x,t)$, its differential

$$\delta f=(\partial f/\partial x)\delta x+(\partial f/\partial t)\delta t$$

subjected to the condition $\delta f=0$ reads $v_x\partial f/\partial x+\partial f/\partial t=0$ with $v_x=\delta x/\delta t$; so with vector notation $\boldsymbol{v}\cdot\nabla f=-\partial f/\partial t$. Putting by definition $\boldsymbol{G}=\boldsymbol{v}f$, where $\boldsymbol{G}$ is an arbitrary vector to be specified, the result $\nabla\cdot\boldsymbol{G}-f\nabla\cdot\boldsymbol{v}=-\partial f/\partial t$ yields $\nabla\cdot\boldsymbol{G}=-\partial f/\partial t$ once having put $\nabla\cdot\boldsymbol{v}=0$. The solenoidal character of the velocity vector excludes sinks or sources of matter crossing from inside or outside the surface of an ideal flux pipe around $v$. Also, it is clearly convenient to identify the arbitrary vector $\boldsymbol{G}$ with the flux vector $\boldsymbol{J}$ and thus $f=C$. If so, then $\nabla\cdot\boldsymbol{J}=-\partial f/\partial t$ yields the component

$$\boldsymbol{J}=C\boldsymbol{v},\quad \frac{\partial}{\partial x}\left(D\frac{\partial C}{\partial x}\right)=\frac{\partial C}{\partial t}$$

*i.e.* the definition of mass flux and the one dimensional second Fick law.

Eventually, Equation (26) reads with the help of Equations (3) and (27) as follows

$$\frac{m}{M_o}\delta E=k_BT\frac{m}{M_o}\delta\log(m/M_o),\quad M_o=const$$

Suppose now that $m$ is the $j$-th mass in a system constituted of a number $j_{tot}$ of masses, *i.e.* actually it is regarded here as $m_j$. Next sum up this equation over $j$, *i.e.* over all masses of the system; one finds thus

$$\delta E=k_BT\sum_j\Pi_j\delta\log\Pi_j,\quad M_o=\sum_j m_j,\quad \Pi_j=\frac{m_j}{M_o}.$$

Since by definition $\delta\log\Pi_j=\log\Pi_j-\log\Pi_o$, assuming $\Pi_o$ independent of the index $j$ one finds

$$\frac{\delta E}{T}=k_B\sum_j\Pi_j\log\Pi_j+S_o,\quad S_o=-\log(\Pi_o).$$

This equation defines the entropy $S$ a function $S_o$ apart as

$$\frac{\delta E}{T}=-S+S_o,\quad S=-k_B\sum_j\Pi_j\log\Pi_j;$$

note that $S_o$ is not necessarily constant, simply it does not depend upon *j*, e.g. it can depend on *T* or pressure and so on. So at constant *T* one finds $\delta E - TS_o = -TS$. These contextual results show that the driving force of the Fick laws is the entropy increase, *i.e.* the second principle of the thermodynamics.

All this is linked to the further information provided by Equation (25). Noting that $v\delta m = \delta(mv) - m\delta v$, this equation yields

$$\frac{\delta E + mD\delta v}{\delta v} = D\frac{\delta(mv)}{\delta v}; \tag{28}$$

this equation relates $\delta E$ to $\delta(mv)$; depending on the sign of this latter, one can have mass fluctuations corresponding to $\delta v$, to which is related the energy fluctuation $\delta E$. If in particular $mv = const$, then $v\delta m = -m\delta v$ implies $h\delta v = Dv\delta m$. It appears that the energy fluctuation $\delta v$ of the radiation is linked to the evaporation or deposition rates $mv$ of matter on or from the contact wall; their relative balance determines the increasing or decreasing amount of mass in the cavity correspondingly to the concurring oscillations of $\delta v$. Eventually note that the left hand side of Equation (28) defines the energy $\varepsilon = hv + mvD$ to which contribute not only the radiation but also the matter through its evaporation rate $mv$. This conclusion automatically includes the interaction between photon and solid matter, without excluding of course that of photons with gas particles evaporated from the surface. Moreover the model provides thermodynamic information able to describe both the equilibrium state of the system and its transient deviation during its fluctuation.

Combine now Equations (24) and (6) to eliminate *v*; as $h = mv\lambda$ owing to Equation (4) and thus

$$E = \gamma pc = \gamma\varepsilon v/c,$$

the result is

$$E\varepsilon = \zeta_0\gamma(pc)^2 + \zeta_1 E^2 pc/2 + (p_o c)^2 w, \quad w = \frac{v}{c\gamma} = \frac{pc}{\gamma\varepsilon}.$$

So

$$\frac{E\varepsilon}{w} = \frac{Z}{w}(pc)^2 + p_o c^2 const + \frac{\zeta_1 E^2 pc}{2w}, \quad Z = \zeta_0\gamma. \tag{29}$$

Put preliminarily $\zeta_1 = \zeta_2 = 0$, *i.e.* neglect the first and second order terms of Equation (23); this equation reduces to $\zeta \approx \zeta_0$ and reads then

$$\varepsilon'^2 = (p'c)^2 + (m'c^2)^2, \quad \varepsilon' = \varepsilon\sqrt{\frac{E}{\varepsilon w}}, \quad p' = p\sqrt{\frac{Z}{w}}, \quad m' = \frac{\sqrt{p_o const}}{c} = m\sqrt{\frac{const}{mc}}. \tag{30}$$

Equations (29) and (30) concern both arbitrary square energies, a scale factor apart for the three quantities characterizing the initial $\varepsilon$ and *p* of Equation (6), and thus are physically equivalent provided that $p' = \varepsilon' v'/c^2$; this requires of course $\varepsilon' v'/p' = \varepsilon v/p$ and implies an appropriate scale factor that converts the initial *m* to $m'$. In effect the variables of the problem are three, *i.e* $v$, *m* and *v*, *i.e. n*; whatever the specific value of $Z = Z(E)$ might be, the constrains of these positions are three as well, *i.e.* the factors linking the last three equations. It is worth noticing that if $m = m'$ for $const = mc$, then $\varepsilon'^2 - (p'c)^2 = \varepsilon^2 - (pc)^2 = (mc^2)^2$ by definition. Since this conclusion holds even if referred in particular to different coordinate systems in reciprocal constant motion, this shows that these equations represent invariant expressions of energy.

Consider now that in Equation (29)

$$\frac{\zeta_1 E^2 pc}{2w} = \frac{\zeta_1}{2} E^2 \gamma\varepsilon = \frac{\zeta_1}{2} \gamma^2 (pc)^2 \varepsilon:$$

this term having the form $p^2\varepsilon$ with $\zeta_1 < 0$ is a well known result of quantum gravity, which solves three cosmological paradoxes [8]. In conclusion, combining the zero order approximation of Equation (23) with Equation (6) one finds the classical expression of relativistic energy; the additional first order term accounts for the quantum correction of the rest energy $mc^2$ of cosmological significance. The Section 7 will show that actually even this result is not accidental.

The fact that $\zeta_1 < 0$ fits the physical meaning of the literature result stimulates a further idea. As $E = (v/c)\varepsilon_0$ according to Equation (3), Equation (23) yields at the first order

$$\zeta = \frac{\zeta_0 c + v\zeta_1\varepsilon_0}{c} = \frac{v'}{c}, \quad v' = \zeta_0 c + v\zeta_1\varepsilon_0;$$

in effect $\zeta_1 < 0$ is compatible with $v' < c$. The form $v' = av + b$ suggests that $v$ and $v'$ could be, at least approximately, velocity components expressed in different reference systems. Is thus attracting the idea of implementing $v'$ to define $n' = c/v'$ and then $\varepsilon'_0 = n'E'$ in analogy with Equation (3) but in a different reference system. Moreover admit for generality that $\varepsilon'_0$ and $E'$ depend on new mass $m'$ and frequency $v'$; thus, differentiating $\varepsilon'_0 = n'E'$ exactly as before to infer Equation (21), one finds

$$n' - \frac{E'}{c} n'^2 \frac{\delta v'}{\delta E'} = \frac{v'}{c}.$$

In effect $v'$ has no peculiarity with respect to $v$ previously introduced; both are arbitrary velocities, both fulfill the same kind of connection between $\varepsilon_0$ and *E*. If this reasoning is correct, then even this result must have a sensible physical meaning. The check is again carried out solving this primed differential equation. One finds

$$v' = \pm \frac{E'c}{\sqrt{E'^2 + c^2 p_{o'}^2}}$$

and thus

$$E' = \pm \frac{v'm'c}{\sqrt{1 - v'^2/c^2}}.$$

Since $E'/c$ is momentum, this result reads

$$p' = \pm \frac{m'v'}{\sqrt{1 - v'^2/c^2}}, \quad \gamma = \sqrt{1 - v'^2/c^2}; \tag{31}$$

*i.e.* is admissible Equation (21) with the right hand side having the form $v/c$ instead of the definition (23) of $\zeta$. In fact, this conclusion is still compatible with Equation (23) itself simply putting $\zeta_0 = 0$ as a particular case in an appropriate reference system. It is instructive to obtain this last result even through a different reasoning.

Calculate via the second Equation (6)

$$\frac{\partial \varepsilon}{\partial v} = -\frac{mc^2}{\gamma^2}\frac{\partial \gamma}{\partial v}.$$

Split this equation putting by definition

$$\frac{\partial \varepsilon}{\partial v} = \frac{p}{\gamma^2}, \quad p = -mc^2 \frac{\partial \gamma}{\partial v}. \tag{32}$$

The second position, allowed in principle by dimensional reasons, allows to handle the first equation as follows with the help of the first Equation (6)

$$\gamma^2 = -p\frac{\partial v}{\partial \varepsilon} = v\frac{\partial p}{\partial \varepsilon} - \frac{\partial (pv)}{\partial \varepsilon}.$$

Note that there is no reference to $\delta v$ in this last result, which instead relates the changes of $p$ and $pv$ to $\mathrm{d}\varepsilon$ of the energy. Assume therefore that these changes are due to $\mathrm{d}m$ and not to $\mathrm{d}v$. In this case, the first Equation (6) yields

$$\gamma^2 = 1 - \frac{v^2}{c^2}. \tag{33}$$

This result is confirmed by the second Equation (32), which yields with the help of the third Equation (6)

$$\frac{mv}{\gamma} = -mc^2 \frac{\partial \gamma}{\partial v};$$

this equation reads $\gamma \mathrm{d}\gamma = -\left(v/c^2\right)\mathrm{d}v$ and is easily integrated. The result is

$$\frac{1}{2}\gamma^2 = const/2 - \frac{1}{2}\frac{v^2}{c^2}, \tag{34}$$

being $const/2$ the integration constant. It appears that putting $const = 1$ the result coincides with that previously found, despite here has been considered the dependence of $p$ and $\varepsilon$ on $v$. It is easy to realize that only the positions (32) allow a consistent calculation of $\gamma$ in Equations (33) and (34); for instance, replacing Equation (32) with the $\gamma\, \partial\varepsilon/\partial v$ and $p = \left(mc^2/\gamma\right)\partial\gamma/\partial v$, in principle also possible because $\gamma$ is dimensionless, would imply inconsistent expressions of $\gamma$. This conclusion agrees with the result of Equation (31).

It appears in conclusion that the term $\zeta_0$ is enough for the purposes of the present model, while it is confirmed that the zero order term of the series (23) accounts for "classical" relativistic results.

Implement then Equation (21) in the simplest form

$$\delta\varepsilon_0 = \zeta_0 \delta E. \tag{35}$$

Comparing with Equation (19), $\zeta_0$ is nothing else but anyone among the possible values of $k/n^2$. Therefore, the third equation of the problem that regards separately $\langle\delta\varepsilon_0\rangle$ and $\langle\delta E\rangle$ reads

$$\langle\delta E\rangle = \frac{k}{n^2}\langle\delta\varepsilon_0\rangle, \quad \langle\delta E\rangle = \zeta_0\langle\delta\varepsilon_0\rangle; \tag{36}$$

the physical meaning of this equation is to consider the averages of all possible $\delta E$ and $\delta\varepsilon_0$ compatible with the given $E$ and $\varepsilon_0$; in effect the arbitrary changes $\delta E$ and $\delta\varepsilon_0$ are independent of the respective $E$ and $\varepsilon_0$, as already remarked.

In conclusion, to the four variables appearing in (17) correspond three Equations (3), (21) and (36); the free parameter $k$ introduced in (18) is a freedom degree of the problem as a function of which are in principle determinable various $E$, $m$ and $v$, *i.e.* $n$.

These results have been hitherto obtained without specific reference to the black body cavity and even regardless of the Planck formula. The next section concerns just this topic.

## 5. The Black Body

To specify the previous results in the case of radiation in a black body cavity of arbitrary volume $V$, it is useful to consider first the Planck law. Noting that this law reads

$$\rho_{Pl} = 4\pi h \frac{N_\nu}{V_\nu}, \quad V_\nu = \frac{c^3}{\nu^3}, \quad N_\nu = \frac{2}{\exp\left(h\nu/k_B T\right) - 1},$$

let us examine the three factors that define $\rho_{Pl}$.

The degeneracy factor 2 of the Bose statistical distribution of photons with the same energy corresponds to the orthogonal polarizations of light [9], to which is due the usual elliptic polarization of a light beam of frequency $\nu$.

The factor $4\pi$ suggests an integration over a solid angle $\mathrm{d}\Omega$. The physical meaning of this statement is clarified below. It is anticipated here that the integral concerns the random impacts of photons on various points of the internal surface of the cavity because of multiple reflections; accordingly any element of this surface thermalizes the radiation trapped inside $V$.

The notation $N_\nu/V_\nu$ of the number density of photons with frequency $\nu$ emphasizes that just the wavelength $c/\nu$ defines the volume $V_\nu$ enclosing a cluster of $V_\nu$ photons with the same frequency $\nu$, whereas instead the true volume $V$ of the cavity is seemingly irrelevant; it is replaced by the local volume defined by the cluster of photons themselves, supposed of course non-interacting at the usual temperatures at which is modeled and tested the black body radiation law. Also this crucial point is concerned below.

With these hints, is really easy to infer the Planck result even in the present physical frame only.

First of all, $N_\nu$ is found implementing once more Equations (26) and (27). Integrate Equation (26) with the help of Equation (27) between two arbitrary energies $\varepsilon$ and $\varepsilon_o'$;

one finds $(\varepsilon-\varepsilon'_o)/k_BT=\log(C/C_o)$. It is possible to write $C=C_o\pm C_\nu$, because in principle $C$ can be greater or lower than the constant $C_o$ of interest for the present reasoning; then this last equation reads $C_o/C_\nu=(C_o/C_\nu)\exp(\Delta\varepsilon'/k_BT)\pm1$. Of course $C_o=m_o/V_o$ and $C_\nu=m_\nu/V_o$ are referred to the same arbitrary volume $V_o$; moreover the masses $m_o$ and $m_\nu$ are proportional to the respective numbers $N_o$ and $N_\nu$ of particles, as they refer to a unique material or kind of particle. Hence one finds with these positions

$$N_\nu=\frac{N_o}{\exp\left(\dfrac{\varepsilon-\varepsilon_o}{k_BT}\right)\pm1},\quad \varepsilon_o=\varepsilon'_o-k_BT\log(C_o/C_\nu) \tag{37}$$

Note that $\varepsilon_o$ has the same form of $\mu$ of Equation (27) a constant $\varepsilon'_o$ apart, *i.e.* it is chemical potential. This result is nothing else but the well known statistical distribution of bosons and fermions as a function of the energy; their occupancy numbers of quantum states are inferable in general from the respective profiles as a function of temperature for either sign, instead of being postulated "a priori".

This point does not need further comments. Here, with the minus sign and putting $\Delta\varepsilon=h\nu$, one calculates the Planck equation.

The number density $N_\nu/V_\nu$ is calculated via a variable volume $V_\nu$ dependent upon the wavelengths allowed in the cavity compatibly with the fixed real volume $V$.

Let the cavity contain $N_\nu$ photons of frequency $\nu$ that define the energy density $\eta=\eta(\nu,T)$ in the physical volume $V_\nu$; then

$$\eta=\sum_\nu\eta_\nu=\sum_\nu\frac{h\nu N_\nu}{V_\nu},\quad V_\nu=\left(\frac{c}{\nu}\right)^3, \tag{38}$$

where clearly $\nu$ is in general anyone of the $\nu_n$ frequencies allowed in the cavity. In this equation, the wavelength is regarded as measure unit to express the size of each $V_\nu$, which in this way results consistent by definition with the existence of standing waves. Moreover the obvious condition

$$V=\sum_\nu V_\nu, \tag{39}$$

is fulfilled because $V$ has not yet been specified. Whatever $V$ might be, the sum over the various $\nu$ can be replaced by that over an arbitrary real number $n$ via the position $\nu=n\nu_o$, being $\nu_o=c/\Delta x$ the lowest frequency allowed in a cavity of size $\Delta x$. If the various $\nu$ are very close each other, then $n$ can be regarded as a continuous variable; if so, the sum can be replaced by an integral between $n_o=1$, in order to include $\nu_o$, and an arbitrary $n_{\max}>1$.

In this case one would find $V\approx V_o\left(1-n_{\max}^{-2}\right)/2$; the notation emphasizes that replacing sum with integral implies a numerical approximation. Despite this result fulfills the obvious requirements of increasing and finite $V$ for $n_{\max}\to\infty$, one would expect $V>V_o$: by definition, indeed, $V_o$ is the volume $(c/\nu_o)^3$ pertinent to the lowest frequency only. Moreover if $n_{\max}$ would be plain real number, the limit $n_{\max}\to1$ would yield $V\to0$; so the energy $\eta V$ inside the cavity should vanish, unless admitting

$\eta \to \infty$. These inconsistencies, not merely numerical but physical, can be due to nothing else but to the low values of $n$ contributing to the sum badly approximated by the integration; indeed it is true that $n_{\max} \to \infty$ behaves in fact like a continuous variable. In effect the contribution of the low values of $n$ is underestimated by the integration. Examine therefore the chance that n can take integer values only: in this case one finds $V > V_o$ even for the lowest $n_{\max} = 2$, whereas $V = \zeta(3) \approx 1.202 V_o$ for $n_{\max} \to \infty$. Also, $V$ remains anyway finite because $n$ cannot longer approach arbitrarily to 1.

It is known in effect that the steady wavelengths $\lambda_n$ allowed within a range $\Delta x$ must fulfill the condition

$$\Delta x = n\lambda_n, \tag{40}$$

with $n$ integer; in other words, the electric field of an electromagnetic wave must vanish at the boundaries of its physical volume of confinement, correspondingly to wave nodes at the boundaries.

The seemingly innocuous position (39) implies thus the energy quantization in the cavity. Equations (2) and (3) yield indeed

$$E = \frac{hc}{\Delta x} = \frac{hc}{n\lambda_n} = \frac{h\nu_n}{n},$$

*i.e.* $h\nu_1 = E$ and $h\nu_2 = 2E$ and so on for all $\lambda_n$ allowed in the cavity once regarding $\Delta x$ as its size. If the photons are assumed non-interacting at the usual $T$ of interest for the black body physics, the sum (38) consists of independent terms. Considering one of these terms, $\eta_\nu$, and differentiating it, one finds at the first order

$$\eta_\nu = \frac{h\nu N_\nu}{V_\nu}, \quad \delta\eta_\nu = \frac{\partial \eta_\nu}{\partial \nu}\delta\nu + \frac{\partial \eta_\nu}{\partial T}\delta T; \tag{41}$$

thus

$$\delta\eta_\nu = h\left(\frac{N_\nu}{V_\nu} + \nu\frac{\partial(N_\nu/V_\nu)}{\partial \nu}\right)\delta\nu + h\frac{\partial(\nu N_\nu/V_\nu)}{\partial T}\delta T. \tag{42}$$

To highlight the physical meaning of the differentials $\delta\nu$ and $\delta T$, implement this equation to calculate the energy density per unit range $\delta\nu$, *i.e.*

$$\rho = \frac{\delta\eta_\nu}{\delta\nu} = h\frac{N_\nu}{V_\nu} + h\nu\frac{\partial(N_\nu/V_\nu)}{\partial \nu} + h\frac{\partial(\nu N_\nu/V_\nu)}{\partial T}\frac{\delta T}{\delta\nu}. \tag{43}$$

All frequencies allowed in the cavity contribute to $\eta$ according to Equation (41), whereas Equation (43) selects some frequencies in the range $\delta\nu$: *i.e.* $\rho$ is an energy density per unit frequency range. All addends share the number density $N_\nu/V$ of a cluster of photons with the same frequency; regard thus this ratio as characteristic property of the cluster. Consider now that the thermal equilibrium inside the cavity requires the exchange of energy between the various $V_\nu$ existing in the cavity. Since however the photons of each cluster have been assumed non-interacting, this exchange cannot be that between different clusters; hence the thermalizing interaction can be nothing else but that with the cavity surface enclosing all photon clusters and possibly with the gas matter evaporated from the walls of the cavity. This fact suggests that the equation

$$\rho' = \int \rho \mathrm{d}\Omega \tag{44}$$

defines the radiation energy density per unit frequency $\rho'$ at the thermal equilibrium with the whole internal surface of the cavity; the corresponding $\rho = \mathrm{d}\rho'/\mathrm{d}\Omega$ represents instead the local interaction of each cluster of photons per unit solid angle, *i.e.* with any elementary surface element $\mathrm{d}s = \Delta r^2 \mathrm{d}\Omega$. In effect the integral represents by definition the sum of all local interactions $\rho \mathrm{d}\Omega$ of the photon cluster with elementary elements $\mathrm{d}s$ of internal surface of the cavity; this supports the idea of regarding $\rho$ as a local quantity and $\rho'$ as an average global quantity. In other words, the integral corresponds to and represents the cumulative effect of all internal reflections of each photon cluster consistent with the physical model of black body cavity. This is equivalent to say that $\rho$ concerns the local thermal equilibrium of the photon cluster with *one* arbitrary surface element $\mathrm{d}s$ only, $\rho'$ represents the complete thermal equilibrium after interaction of the cluster with the *whole* surface of the cavity. So $\rho$ and $\rho'$ differ numerically because of the amount of corresponding energy density exchanged between radiation and surface.

If Equation (44) leads to the correct formulation of the Planck law, then it also proofs indirectly that the photon thermalization mechanism occurs at the surface of the cavity.

The integration of $\int \rho \mathrm{d}\Omega$ is immediate admitting that the interaction process is isotropic, *i.e.* the energy exchange occurs uniformly for all frequencies and that any allowed $\nu$ is not appreciably perturbed by the small energy loss; being the radiation field at the equilibrium uniformly distributed inside the cavity, there is no dependence of $\nu$ upon the arbitrary direction along which is defined $\delta\Omega$. So the result of the integration is simply $\rho' = 4\pi\rho$. Equation (43) yields therefore the following energy density per unit frequency thermalized by all possible paths of the ν-th cluster of photons in the cavity:

$$\rho' = 4\pi h \frac{N_\nu}{V_\nu} + 4\pi h\nu \frac{\partial\left(N_\nu/V_\nu\right)}{\partial\nu} + 4\pi h \frac{\partial\left(\nu N_\nu/V_\nu\right)}{\partial T}\frac{\delta T}{\delta\nu}. \tag{45}$$

Noting that

$$\frac{\partial\left(\nu N_\nu/V_\nu\right)}{\partial T} = \frac{\partial\left(\nu N_\nu/V_\nu\right)}{\partial\nu}\frac{\partial\nu}{\partial T} = \frac{N_\nu}{V_\nu}\frac{\partial\nu}{\partial T} + \nu\frac{\partial\left(N_\nu/V_\nu\right)}{\partial\nu}\frac{\partial\nu}{\partial T},$$

Equation (45) reads then

$$\rho' = 4\pi h \frac{N_\nu}{V_\nu} + 4\pi h\nu \frac{\partial\left(N_\nu/V_\nu\right)}{\partial\nu}\left(1 + \frac{\partial\nu}{\partial T}\frac{\delta T}{\delta\nu}\right) + 4\pi h \frac{N_\nu}{V_\nu}\frac{\partial\nu}{\partial T}\frac{\delta T}{\delta\nu}. \tag{46}$$

This expression can be considerably simplified because

$$\frac{N_\nu}{V_\nu} = N_o c^{-3}\frac{\nu^3}{\exp\left(h\nu/k_B T\right) - 1}$$

has a maximum as a function of $\nu$, which suggests $\partial N_\nu/\partial V_\nu = 0$ for an appropriate value $\nu^*$ of the $\nu$-th frequency. In this case the second addend of Equation (45) vanishes for this particular value $\nu = \nu^*$. Hence, replacing $N_o = 2$ in Equation (37) to account for the light polarization states with the same $\nu$ in the Bose function and calculating Equation (46) with $\nu = \nu^*$, the result at $T = const$ is simply

$$\rho_{Pl} = 4\pi h \frac{\nu^{*3}}{c^3} \frac{2}{\exp\left(h\nu^*/k_B T\right)-1} \left.\frac{\partial\left(N_\nu/V_\nu\right)}{\partial \nu}\right|_{\nu=\nu^*} = 0. \tag{47}$$

Therefore, the plain Planck law corresponds to the particular set of frequencies that, among the ones allowed in the cavity, maximize the number density of photons with a given energy at a fixed *T*.

Actually, however, no physical reason requires $\nu = \nu^*$ and $T$ really constant; nevertheless, the analytical form of the first addend of Equation (46) is identical to that of $\rho_{Pl}$ in Equation (47). This suggests that $\rho_{Pl}$ is still given in general by the first addend of Equation (46) even though calculated with a frequency $\nu \neq \nu^*$ and thus without the constrain on $T$ that annul the other terms; these terms account therefore for the frequency and temperature fluctuations with respect to the zero order term represented by the Planck function. This conclusion clarifies that $\delta\nu$ and $\delta T$ represent just the frequency and temperature fluctuations of the cavity.

In the present model it appears therefore that:

- The interaction between degenerate photon clusters and internal walls of the cavity is responsible for the thermalization mechanism.
- The fluctuations are inferred contextually to the Planck law itself.

To emphasize these points, it is necessary now to link these fluctuations with Equations (19) and (20). As expected, the fluctuation is given by temperature and frequency deviations of $\rho'$ with respect to the mere equilibrium Planck term; simple considerations show indeed that the fluctuation terms can have in principle positive or negative sign.

Now it is possible to tackle the problem of describing the cavity for $\nu \neq \nu^*$ and $\delta T \neq 0$, *i.e.* when both frequency and temperature are allowed to fluctuate.

## 6. Black Body Fluctuation

The result (25) and Equation (44) imply the involvement of the material constituting the wall of the cavity to reach the condition of thermodynamic equilibrium of photons therein confined. In particular $\delta E$ appears related to $\delta m$ and $\delta(mv)$, *i.e.* to the material evaporating from the internal surface of the shell and present in the cavity together with the radiation field. This is confirmed by the mean square velocity $\langle v^2 \rangle$ of matter particles present in gas phase in the cavity contextually inferred. It is known from the elementary kinetic theory of gases that $\langle v^2 \rangle$ is related to $k_B T$. Even though the photons are admitted non-interacting, their thermalization process occurs by interaction both with the internal wall of the cavity and with the amount of matter expectedly evaporated and trapped in the cavity together with the radiation itself; clearly the gas phase is at the thermal equilibrium with the cavity wall.

For sake of clarity, collect together Equations (3), (27) and (25); one finds

$$\frac{h}{n}=h\frac{h\nu}{mc^2}=\frac{mD}{k}=\frac{m\beta k_BT}{k}=\frac{k_BT}{k\nu}. \tag{48}$$

These equations evidence in particular

$$k\frac{h\nu}{n}=k_BT \quad i.e. \quad k\frac{\varepsilon_0}{n^2}=k_BT, \tag{49}$$

whereas Equation (19) reads

$$kE^2=\varepsilon_0^2\frac{\delta E}{\delta\varepsilon_0}. \tag{50}$$

Since $k$ has been defined as a mean value in Equation (18), let then be

$$k=q\langle n^2\rangle,$$

being $q$ an arbitrary constant. Then, Equation (36) yields

$$\frac{q\langle n^2\rangle}{n^2}=\frac{\langle\delta E\rangle}{\langle\delta\varepsilon_0\rangle}.$$

Since Equation (49) reads

$$\varepsilon_0=\frac{n^2}{q\langle n^2\rangle}k_BT,$$

so that merging these equations one finds

$$\varepsilon_0=\frac{\langle\delta\varepsilon_0\rangle}{\langle\delta E\rangle}k_BT, \tag{51}$$

the result obtained via Equation (50) is

$$q\langle n^2\rangle E^2=\frac{\langle\delta\varepsilon_0\rangle^2}{\langle\delta E\rangle^2}(k_BT)^2\frac{\delta E}{\delta\left(k_BT\,\delta\langle\varepsilon_0\rangle/\langle\delta E\rangle\right)}=\frac{\langle\delta\varepsilon_0\rangle}{\langle\delta E\rangle}(k_BT)^2\frac{\delta E}{\delta(k_BT)}.$$

Therefore

$$q\langle n^2\rangle E^2=\frac{\langle\delta\varepsilon_0\rangle}{\langle\delta E\rangle}(k_BT)^2\frac{\delta E}{\delta(k_BT)}$$

yields

$$U^2=(k_BT)^2\frac{\langle\delta\varepsilon_0\rangle}{q\langle\delta E\rangle}\frac{\delta E}{\delta(k_BT)}, \quad U^2=\langle n^2\rangle E^2; \tag{52}$$

hence

$$\langle U^2\rangle=(k_BT)^2\frac{\langle\delta\varepsilon_0\rangle}{q\langle\delta E\rangle}\frac{\langle\delta E\rangle}{\delta(k_BT)}=(k_BT)^2\frac{\langle\delta\varepsilon_0'\rangle}{\delta(k_BT)}, \quad \varepsilon_0'=\frac{\varepsilon_0}{q}. \tag{53}$$

As $\langle\delta\varepsilon_0'\rangle=\delta\langle\varepsilon_0'\rangle$, this is just the famous Einstein equation [10]. To find this result, Einstein quoted the energy of a sub-volume enclosed by a large volume, both concurring to the total volume of the cavity and exchanging energy. Here the role of the smaller volume is proportional to $\varepsilon_0$, whose fluctuation is the source function of the Einstein model of a closed system. Actually this appears in Equation (44), because the photons are thermalized just impacting against the wall of the cavity, which is therefore the effective source of the photon energy. So it appears clearly that the fluctuations are controlled by the matter constituting the walls of the cavity; this conclusion has been in effect assumed in the paper [5]. If in fact $\delta\varepsilon_0 = 0$, then $\langle U^2 \rangle = 0$: *i.e.* the average energy is controlled by the matter at the walls of the cavity. Consistently $U^2$ is also related itself to the radiation field $E^2$, as it appears in Equation (52). Clearly it is reasonable to put here

$$\delta E = h\left(\nu - \nu^*\right).$$

Replacing Equations (51) into (53), one finds

$$\left\langle U^2 \right\rangle = \left(k_B T\right)\varepsilon_0 \frac{\langle \delta E \rangle}{\delta\left(k_B T\right)}, \quad \delta E = \nu - \nu^*. \tag{54}$$

According to the previous considerations, $\langle U^2 \rangle = 0$ for $\nu = \nu^*$; this confirms that the left hand side of Equation (54) is a fluctuation energy. Equation (51) yields then the relationship between frequency and temperature fluctuations

$$\langle \delta E \rangle = \left\langle \delta\left(k_B T\right)\right\rangle = \delta\left\langle k_B T \right\rangle \quad i.e. \quad \langle \delta\nu \rangle = \frac{k_B}{h}\langle \delta T \rangle. \tag{55}$$

## 7. Discussion

The fluctuations are likely the most typical manifestation of the probabilistic character of the quantum world, while also being the most striking evidence of the quantum uncertainty. Nevertheless, elementary and straightforward considerations have shown that the equations describing the fluctuations are also compliant with relativistic corollaries: both have been concurrently inferred from Equation (1) in a unique theoretical frame. Despite the deterministic character of the relativity, the results so far outlined emphasize this seemingly surprising connection. Actually a similar conclusion was already found also in [7] implementing an operative definition of space time, *i.e.* introducing *ab initio* the quantity $hG/c^2$ as a basic postulate to be handled subsequently likewise any fundamental physical law.

First of all, the present model plugs the problem of the black body radiation and its fluctuations in a wide context of physical laws having prospective interest for the non-equilibrium physics. The quantum basis of the Fick law is important because various physical properties, e.g. the heat and electrical conductivities, have analogous form; here, in particular, the diffusion equations are in principle necessary to account for the unstable concentration gradients reasonably expected in gas phase due to random concentration fluctuations of the matter evaporated from the internal surface of the cavity. In effect the dynamics of matter particles that diffuse from the walls of the cavity contributes to the thermalization process; in this respect, the model introduces concurrently even the free energy and entropy concepts useful to infer the Clausius-Clapeyron equation governing the vapor pressure and thus the amount of matter in gas phase filling the cavity together with the radiation. In view of that, the Planck law has been inferred in order to involve since the beginning the solid matter confining the photons and even their energy quantization and statistical distribution law. The interaction of photons with matter appears in fact essential to justify the thermalization mechanism. Strictly speaking, the radiation with wavelength larger than the finite size cavity should not be consistent with the standard approach to the Planck law; here however this problem is bypassed since the cavity volume $V$ is not predetermined, rather it is determined by the radiation wavelengths themselves via the terms (39). Thus it is by definition compliant with the arbitrary size $\Delta x$ defining the allowed frequencies according to Equations (1), (2) and (38). For these reasons, is reductive the model [10] focused on the radiation field in the cavity only.

The black body radiation field and its fluctuations have been contextually inferred merging two separate paths: the one from Equations (14) to (20) is apparently independent on that leading from Equation (45) to Equation (53). The former series of equations does not refer specifically to the black body radiation, it introduces relationships between changes of dynamical variables that hold in general. The latter series of equations describes specifically the black body radiation under the boundary condition of Equation (20), which also implies Equations (21) to (24); this second path links the frequency and mass fluctuations, in agreement with Equations (4) to (9). Then, Equation (36) introduces the thermal equilibrium of Equation (50) leading to Equation (53).

Yet other significant results are also easily inferable from the previous considerations of the Section 4.

For example, combining Equations (26) and (27) with Equation (28) one finds at constant $T$

$$D\frac{\delta(m\nu)}{\delta\nu}=-\frac{\delta\beta}{\delta\nu}\frac{D}{\beta^2}=-\frac{k_BT}{\beta}\frac{\delta\beta}{\delta\nu}=-k_BT\frac{\delta\log\beta}{\delta\nu}.$$

The equation $D\delta(m\nu)=-k_BT\delta\log\beta$ is easily integrated; calling $\beta_0$ the integration constant, the solution

$$-\frac{\int D\delta(m\nu)}{k_BT}=\log\beta-\log\beta_0$$

yields

$$\frac{\beta}{\beta_0}=\exp\left(-\frac{\int D\delta(m\nu)}{k_BT}\right).$$

Owing to the first Equation (27) put then

$$\frac{D/k_BT}{D_0/k_BT}=\exp\left(-\frac{\int D\delta(m\nu)}{k_BT}\right),\quad D_0=\beta_0k_BT,$$

being $D_0$ a constant diffusion coefficient corresponding to the integration constant

$\beta_0$. The conclusion is

$$D = D_0 \exp\left(-\frac{U_{act}}{k_B T}\right), \quad U_{act} = \int D\delta(mv)$$

This Arrhenius-like equation is a well known property of the diffusion coefficient, whose quantum origin introduces the activation energy as a consequence.

Other important equations of processes activated by the temperature follow this kind of dependence upon $k_B T$.

A further significant result is obtained from Equation (6), assuming that the momentum $p$ is time dependent variable. This compels regarding the wavelength $\lambda$ as time variable itself, as in effect it is possible because no restrictive hypothesis has been introduced about $p$ and thus about $\lambda$ in Equations (2) and (4). Deriving thus $p = h/\lambda$ with respect to time in the reference system $R$ previously introduced to define $\Delta t$ and $\Delta x$ of Equation (2), one finds

$$\dot{p} = -\frac{h}{\lambda^2}\dot{\lambda}, \quad h\dot{\lambda} = mc\lambda_C\dot{\lambda}, \quad \dot{\lambda} = \frac{\delta\lambda}{\delta\Delta t}. \tag{56}$$

It is possible to expand in series $\dot{\lambda}$ around an arbitrary constant value $\dot{\lambda}_o$, e.g.

$$\dot{\lambda} = \dot{\lambda}_o + \sum_j a_j\left(\lambda - \lambda_o\right)^j, \tag{57}$$

being $a_j$ appropriate coefficients. Implement Equation (40) to express again length $\Delta x$ as a function of wavelength $\lambda$; here, however, $\lambda$ is the momentum wavelength of Equation (5). To highlight the physical meaning of the series expansion, retain preliminarily the constant term only and consider two chances of rewriting the first Equation (56). Eliminate $h$ from Equation (56), replacing it via the Planck mass $m_{Pl} = \sqrt{\hbar c/G}$ and fine structure constant $\alpha = e^2/\hbar c$; so, being by definition

$$h = 2\pi m_{Pl}^2 G/c = 2\pi e^2/\alpha c, \tag{58}$$

Equation (56) reads at the zero order of approximation of the series expansion according to Equation (40)

$$\dot{p} = F \approx -\frac{2\pi\dot{\lambda}_o}{c} G\frac{n^2 m_{Pl}^2}{\Delta x^2} = -\frac{2\pi\dot{\lambda}_o}{c}\frac{1}{\alpha}\frac{n^2 e^2}{\Delta x^2}, \quad \Delta x = n\lambda. \tag{59}$$

Since $F$ is actually the component of a force along $\Delta x$, which can have both signs, consider for brevity of notation its absolute value only. This expression reads then

$$|F| \approx G\frac{m'm''}{\Delta x^2} = \frac{e'e''}{\Delta x^2}, \tag{60}$$

having put

$$e' = en'\sqrt{\frac{2\pi\dot{\lambda}_o'}{\alpha c}}, \quad e'' = en''\sqrt{\frac{2\pi\dot{\lambda}_o''}{\alpha c}},$$

$$m' = m_{Pl} n'\sqrt{\frac{2\pi\dot{\lambda}_o'}{c}}, \quad m'' = m_{Pl} n''\sqrt{\frac{2\pi\dot{\lambda}_o''}{c}}, \quad n'n''\sqrt{\dot{\lambda}_o'\dot{\lambda}_o''} = n^2\dot{\lambda}_o.$$

As $n$ and $\dot{\lambda}_o$ are arbitrary, likewise the primed and double primed quantities, these approximate definitions of force correspond to the Newton and Coulomb interactions between arbitrary masses and steady charges, which have thus analogous form. In effect, at this level of approximation, neither term of the equality (60) depends on $v$. The only term that could introduce charge velocity is $\dot{\lambda}_o$; yet no reason requires assuming $\dot{\lambda}_o(v)$. In lack of specific hypotheses, therefore, $\dot{\lambda}_o$ is consistent with stationary charges. In the case of gravity force, it is well known that the Newton law is generalized by the general relativity; it is reasonable therefore to expect that the terms of series expansions, here neglected preliminarily, account for the necessary relativistic corrections of the plain Newton law. To demonstrate this statement, calculate via Equation (56) the potential energy $U$ generated by the mass $m$ defining $p$ at any point $\Delta x_r = n_r \lambda_r$; the subscript emphasizes that $\lambda_r$ concerns the radial distance of $m''$ from $m'$. As by definition $F = -\partial U/\partial \lambda_r$, Equations (56) and (57) yield

$$U = -h\frac{\dot{\lambda}_o}{\lambda_r} - \Gamma, \quad \Gamma = h\int_{\lambda_r}^{\infty}\sum_j a_j\left(\lambda - \lambda_o\right)^j \lambda^{-2}\mathrm{d}\lambda. \tag{61}$$

Whatever the value of $\Gamma$ might be, depending on the series coefficients $a_j$, is remarkable the fact that the potential here inferred differs from the Newtonian form just because of the presence of terms neglected in the classical Equations (59) and (60). It is well known that the perihelion precession of orbiting bodies is correctly calculated in the general relativity by potential terms additional to the mere $-Gm/r$, which however cannot be justified in the plain Newton model [11]. Here, in effect, additional terms appear as a natural consequence of $\dot{\lambda}$ of Equation (56): there is no reason to assume that $\dot{\lambda}$ be equal to the constant $\dot{\lambda}_o$ only, being instead reasonably expectable a more general form like that of Equation (57). Actually it is easy to show that $U$ cannot be equal uniquely to the first addend; owing to Equation (60) one would infer indeed

$$\dot{\lambda}_o = \frac{m'm''}{m_{Pl}^2}\frac{c}{2\pi n_r}, \quad \Delta x = n_r \lambda_r \tag{62}$$

in the mere Newtonian approximation of Equation (59). If so, however, the ratios $m'/m_{Pl}$ and $m''/m_{PL}$ in principle arbitrary likewise $n_r$, could admit $\dot{\lambda}_o > c$, which is however impossible. So a correction term, *i.e.* $\Gamma$, is necessary to define $U$ by comparing Equations (56) and (61). This conclusion confirms therefore that the terms of the sum (57), neglected for simplicity in Equation (60), are in fact essential to agree with the finite light speed and have thus relativistic valence.

It is possible to show the validity of these conclusions, which should hold for the Coulomb law as well, by demonstrating how to find well known results of the general relativity as a consequence of Equation (19).

To this purpose it is necessary to generalize what $\delta m$ actually means in $\delta\varepsilon_0 = c^2\delta m$. The previous considerations about the black body cavity have emphasized that $\delta m$ concerns the material evaporated from the internal walls of the cavity; as the temperature fluctuation modifies the vapor pressure of the cavity material, $\delta m$ refers to the change of amount of material evaporating from or condensing on the internal wall of the cavity according to the Clausius-Clapeyron equation. However the mass fluctuation can have a more general meaning, directly related to the concept of energy quantum fluctuation $\delta\varepsilon$ itself; e.g. it is possible in principle that $\varepsilon_0 + \delta\varepsilon = (m + \delta m)c^2$. Of course this effect is reasonably negligible in the case of a black body cavity, owing to the small $\delta m$ related to the usual cavity temperature and its fluctuations. Yet the following examples aim to show that $\delta m = (m' + m'') - m$ is conceptually significant in principle, even though $|\delta m| \approx 0$. If such a mass fluctuation is indeed allowed to occur, then it is significant to investigate the behavior of the transient formation of $m'$ and $m''$ related to $\delta m$. Expectedly, this brings to the classical Kepler problem, where either mass orbits in the gravity field of the other; so, let us regard for example $m' = m_{or}$ and $m'' = m_{gf}$, where the subscripts stand for *orbiting* and *gravity field*. It is known that the general relativity predicts in this respect two effects, the perihelion precession of $m'$ around $m''$ and the emission of gravitational waves. Since these effects are concomitant, being both features of any orbiting system, the following discussion aims to examine jointly both of them.

Consider first just Equation (19) used to calculate the quantum fluctuations and note that the ratio at right hand side can be rewritten defining $k$ such that $kE^2 = (F_{Pl}\Delta r)^2$: *i.e.*, likewise as done to infer Equations (12) and (59), the definition of Planck force is again implemented here to introduce $G$ into the present problem. So, thanks to the arbitrary numerical factor $k$, the energy $\sqrt{k}E$ is rewritten in order to introduce the arbitrary displacement $\Delta r$ too. Equation (19) reads thus

$$\frac{\delta\varepsilon_0}{\delta E} = \frac{1}{k}\frac{\varepsilon_0^2}{E^2} = \xi^2, \quad \xi = \frac{m''G}{c^2\Delta r}. \tag{63}$$

Introduce now at the right hand side the further mass $m'$; the last equation turns into

$$\frac{\delta\varepsilon_0}{\delta E} = \left(\frac{m'm''G}{m'c^2\Delta r}\right)^2.$$

This result becomes next more familiar via a formal and elementary manipulation. Eliminate $\Delta r$ introducing the modulus of classical angular momentum $|\boldsymbol{M}| = wm'v_{or}\Delta r$ of $m'$, being $v_{or}$ the average orbital velocity of the mobile mass $m'$ and $w$ the numerical coefficient taking into account the vector nature of $\boldsymbol{M}$ and $\boldsymbol{v}_{or}$; this allows considering at the right hand side the modulus of $\Delta\boldsymbol{r} \times \boldsymbol{v}_{or}$. Eliminating thus $m'\Delta r$ at the right hand side, the last formula reads

$$\frac{\delta\varepsilon_0}{\delta E} = w^2\left(\frac{m'm''Gv_{or}}{c^2M}\right)^2, \quad M = wm'v_{or}\Delta r, \quad w \le 1.$$

Eventually, recalling that $v_{or} = c/n_{or}$ according to the initial definition (3), this equation reads

$$\frac{1}{\theta^2}\frac{\delta\varepsilon_0}{\delta E} = \left(\frac{m'm''G}{cM}\right)^2, \quad n_{or} \ge 1, \quad \theta^2 = \frac{w^2}{n_{or}^2} \le 1.$$

At the left hand side, the energies appear through a numerical coefficient times a ratio of the respective fluctuations. Consider now the case where this factor is $\le 1$. Of course, nothing requires just this condition, which however is in principle possible and deserves attention for the related consequence of the initial Equation (63): this particular case is interesting because the left hand side can be regarded as a square probability $\Pi^2$. Hence, it is appropriate to identify $\Pi^2$ with the product of two probabilities, that of angular displacement $\delta\varphi/2\pi$ and that concerning the related tangential velocity component of $m'$; clearly the velocity component of $\boldsymbol{v}_{or}$ corresponding just to the advancement direction $\delta\varphi$ in any point along the orbit has probability 1/3, because the independent local components of $\boldsymbol{v}_{or}$ are three. If so, then $\delta\varepsilon_0/\theta^2\delta E$ takes the meaning compatible with a known formula of the general relativity: indeed, regarding the right hand side of the last result in a probabilistic way as

$$\Pi^2 = \frac{\delta\varphi}{2\pi}\frac{1}{3} = \left(\frac{m'm''G}{cM}\right)^2, \tag{64}$$

one recognizes the well known formula of the perihelion precession. This identification needs however a detailed justification and explanation: helps to this purpose a further result related to the energy loss via gravitational waves, still implied by Equation (63).

It is known that an isolated orbiting system irradiates energy all around in the space; the energy loss causes the orbit shrinking closer and closer towards the central mass. The starting input to demonstrate this effect in the present context is still Equation (19), rewritten identically via Equation (35) as follows

$$\varepsilon_k^2 = k'\zeta_0 E^2, \quad \varepsilon_k = k''\varepsilon_0, \quad k = \frac{k'}{k''^2}; \tag{65}$$

*i.e.* $k$, whatever its specific value might be, has been split into $k'$ and $k''$ suitable to obtain a new value of energy $\varepsilon_k$ from the early $\varepsilon_0$. This is in principle possible because the values of these latter are both arbitrary. The fact that $E = hc/(c/\nu) = hc/\Delta r$, in agreement with the position (10), suggests assuming $k'\zeta_0$ in order that

$$k'\zeta_0 E^2 = \left(G\frac{m'm''}{\Delta r}\right)^2, \quad k' = \frac{1}{\zeta_0}\left(G\frac{m'm''}{hc}\right)^2;$$

moreover if $k'' \propto q/m''$, being $q$ arbitrary proportionality constant, then

$$\varepsilon_k^2 = (qc)^2, \quad k' = \frac{1}{\zeta_0}\left(G\frac{m'q}{k''hc}\right)^2 = k''^2 k, \quad k = \frac{1}{\zeta_0}\left(G\frac{m'q}{k''^2 hc}\right)^2.$$

Hence the first Equation (65) reads

$$\left(G\frac{m'm''}{\Delta r}\right)^2 = (qc)^2,$$

where the right hand side is constant. Integrating now both sides over the solid angle $\mathrm{d}\Omega$, one finds

$$\int\left(G\frac{m'm''}{\Delta r}\right)^2 \mathrm{d}\Omega = const.$$

The square energy at the right hand side is constant; since it consists of fundamental constants only, thanks to the position assumed for $k''$, it is reasonable to put $const = hW_{Pl}$, being $W_{Pl}$ the Planck power. Moreover, if the space around the orbiting system is homogeneous and isotropic, so that the orbiting system irradiates energy uniformly to all directions, one finds

$$4\pi\varepsilon_{\Delta r}^2 = hW_{Pl}, \quad \varepsilon_{\Delta r} = -G\frac{m'm''}{\Delta r}. \tag{66}$$

At this point the quantum uncertainty is of valuable help; it requires that $\Delta r$ and the momentum component $\Delta p_{\Delta r}$ fulfill the condition $nh = \Delta p_{\Delta r}\Delta r$. In the present case, introducing the reduced mass $\mu$ of the orbiting system, the resulting uncertainty equations read

$$\mu\Delta\dot{r}\Delta r = \frac{4\pi n}{W_{Pl}}\varepsilon_{\Delta r}^2, \quad \Delta p_{\Delta r} = \mu\Delta\dot{r} = \frac{nh}{\Delta r}, \quad \Delta\dot{r} = \frac{\delta\Delta r}{\delta\Delta t}, \quad \mu = \frac{m'm''}{m'+m''}. \tag{67}$$

Replacing in Equation (66) $h$ from the second equation, one infers $\Delta\dot{r} = 4\pi n\varepsilon_{\Delta r}^2 / \mu W_{Pl}\Delta r$, *i.e.*

$$\Delta\dot{r} = 4\pi n\frac{G}{\mu c^5\Delta r}\left(G\frac{m'm''}{\Delta r}\right)^2 = 4\pi n\frac{G^3}{c^5\Delta r^3}m'm''\left(m'+m''\right). \tag{68}$$

With the minus sign and $n = 1$, this expression is nothing else but the well known Einstein result of orbit contraction contextual to the emission of gravitational waves: indeed $4\pi$ approximates well the numerical value 64/5 of his original formula. This means that the possible time evolution of the orbiting system described by this energy equation is due to the integer $n$ which can take different discrete values at various times; correspondingly, the orbiting system changes energy and orbital distance from the central mass as well simply according to $n$. This quantum behavior already found in [12] is not surprising, since the starting point of the present reasoning was the quantum law governing the energy fluctuations. The related energy change $\dot{\varepsilon}_{\Delta r}$ of the orbiting system is immediately calculated. It is enough to note in this respect that

$$\frac{\dot{\varepsilon}_{\Delta r}}{\Delta\dot{r}} = \frac{\delta\varepsilon_{\Delta r}}{\delta\Delta r} = F, \quad \dot{\varepsilon}_{\Delta r} = \frac{\delta\varepsilon_{\Delta r}}{\delta\Delta t},$$

where $F$ is force. Simply considering the elementary positions

$$\varepsilon_{\Delta r} = -G\frac{m'm''}{\Delta r} = \frac{\mu\omega^2\Delta r^2}{2}, \quad F = \mu\omega^2\Delta r,$$

one finds

$$\begin{aligned} &\dot{\varepsilon}_{\Delta r} = \Delta\dot{r}\frac{\varepsilon_{\Delta r}}{\Delta r} = \frac{\mu\omega^2\Delta r}{2}\frac{4\pi nG^3}{c^5\Delta r^3}m'm''\left(m'+m''\right) = 2\pi nG\mu^2\frac{\omega^6\Delta r^4}{c^5}, \\ &\omega^2 = \frac{\left(m'+m''\right)G}{\Delta r^3}. \end{aligned} \tag{69}$$

The second equation is well known in the elementary Kepler problem identifying $\Delta r$ with the major semi-axis of the elliptic orbit. This result shows that the orbit size is subjected to change, concurrently to its angular displacement previously introduced;

indeed Equation (69) concerns in particular the perihelion distance. Once again appears the Einstein formula for $n=1$ and without minus sign. The explanation of these two results and their connection with Equation (64) is really simple.

Here $\Delta\dot{\varepsilon}_{\Delta r}$ with $n=1$ represents the energy gap $\varepsilon_{\Delta r(n=2)}-\varepsilon_{\Delta r(n=1)}$ related to the jump of $m'$ between orbits where $\Delta r=\Delta r(n=1)$ to $\Delta r=\Delta r(n=2)$, or in general to any $\Delta r=\Delta r(n>1)$, during the time range $\delta\Delta t$; of course this latter is not a differential d*t*, physically meaningless, but a finite time change of $\Delta t$ necessary for any physical process to occur. This quantum point of view leading to Equations (68) and (69) is coherent with Equation (19) leading to the fluctuation Equation (53). The quantum standpoint also explains the lack of explicit minus sign in both Equations (68) and (69). In the Einstein result the orbital motion progressively decays towards distances closer and closer around the central gravitational mass with gradual energy loss only; accordingly, any orbiting system is destined to merge soon or later its bodies into a unique celestial body. In the present model instead Equation (68) is the distance gap between two contiguous orbits allowed with $\Delta n=1$, *i.e.* $m'$ can in principle decay or be excited towards a lower or higher *n*-th orbits. This also means that two gravitational systems can even exchange "resonant" energy, e.g. by exchanging gravitons, likewise as two atoms of the same kind do by exchanging photons if either of them is in any electron excited state and the other in the fundamental state. It is also evident the analogy with the electrons that do not fall on the nucleus, but occupy stable quantum levels. So the lack of minus sign means that the formulas concern the amount of quantum energy exchanged regardless of whether this energy is released or absorbed by a given orbital system.

Consider now any point of the ellipse at a given time $\Delta t$ and at later time $\Delta t'$; e.g. this point could be, but not necessarily must be, the perihelion. Equation (68) of $\Delta\dot{r}$ shows that this point moves radially from its initial position, as it is evident in the momentum/position uncertainty Equation (67) implementing radial conjugate dynamical variables. Equation (64) accounts instead for the tangential motion of $m'$ in any given point along to the orbit: of course nothing, apart from the algebraic elaboration of the formulas, compels tangential displacement only or radial displacement only of the orbit of $m'$. So, as previously emphasized, Equation (64) on the one hand and Equations (68) and (69) on the other hand simply complete each other in describing the dynamics of a unique phenomenon, *i.e.* the radial and tangential displacements of $m'$ along its orbit that rotates and deforms as a function of time; this is coherent with the fundamental idea of deformation of the space time in the presence of a gravitational mass $m''$. This is in effect the physical meaning of $\dot{\lambda}$ in Equation (56), being $\lambda$ linked to $\Delta x$ via Equations (2) and (4). Note that $m'$ and $m''$ can be exchanged while leaving identical the results: as nothing distinguishes the specific role of either of them from a physical point of view, one concludes that the concepts of gravitational and inertial mass are physically indistinguishable.

The idea of introducing Planck units is fruitful and general, as it is confirmed also in the following reasoning.

Rewrite $\varepsilon_0 = nh\nu$ as

$$\varepsilon_0 = h\nu_m, \quad \nu_m = \nu c/v$$

thanks to Equation (3). So, owing to Equation (21), Equations (19) and (15) yield

$$\frac{\delta\varepsilon_0}{\delta E} = \frac{1}{k}\frac{\varepsilon_0^2}{E^2} = \frac{v_\varepsilon^2}{kv^2}$$

while being owing to Equation (21)

$$\frac{\delta\varepsilon_0}{\delta E} = n - \frac{n^2 E}{c}\frac{\delta v}{\delta E} = n\left(1 - \frac{\varepsilon_0}{c}\frac{\delta v}{\delta E}\right),$$

whence

$$\frac{\nu_m^2}{\nu_0^2} = 1 - \frac{nE}{c}\frac{\delta v}{\delta E} = 1 - \frac{\varepsilon_0}{c}\frac{\delta v}{\delta E}, \quad \nu_0 = \nu\sqrt{kn}. \tag{70}$$

Since $\delta E/\delta(v/c)$ has physical dimensions of an energy, it is possible to put

$$\frac{\delta E}{\delta(v/c)} = F\Delta x' = \frac{F_{Pl}}{2}\Delta x$$

where $F_{Pl} = c^4/G$ is the Planck force; being $\Delta x'$ and $\Delta x$ arbitrary lengths, the energy at the left hand side can be certainly expressed as $F\Delta x'$ and in turn this latter as $F_{Pl}\Delta x/2$. These positions merely implement the general definitions of force and energy. The reason of having introduced the factor 1/2 appears soon after replacing in Equation (70); one finds

$$\frac{\nu_m^2}{\nu_0^2} = 1 - \frac{2mG}{c^2\Delta x}.$$

Note that $\nu_m = \nu_0$ for $m = 0$ and for $\Delta x \to \infty$, *i.e.* in the absence of gravity field; also, owing to Equation (12), $\nu_m \le \nu_0$ implies $\Delta x_{co}/\Delta x < 1$. Just this is the reason of having introduced the factor 1/2: to describe the red shift of a photon moving away from a gravitational mass, the photon must be outside the boundary of its confinement radius (12) of *m*; *i.e.*, the previous limits hold outside the "event horizon" of *m*, otherwise the photon could not freely escape to infinity. In conclusion the last equation reads

$$\frac{\nu_m}{\nu_0} = \sqrt{1 - \frac{2m_r G}{c^2\Delta x}}.$$

Is really significant the fact that also this result of the general relativity is obtained implementing Equations (19) and (18), from which have been obtained Equation (50) and then the black body fluctuation Equation (53). A wider landscape of results of the general relativity is inferred via an "*ab initio*" theoretical model in [7].

On the one hand, the result (60) highlights the quantum origin of the gravity force, simply inferable admitting time dependence of De Broglie momentum wavelength. In this respect Equation (59) prospects an interesting consequence as it yields

$$\frac{1}{A}\frac{G\alpha}{e} = \frac{1}{A}\frac{e}{m_{Pl}^2}, \quad \frac{1}{A} = \frac{2\pi\dot{\lambda}_o}{c}\frac{n^2}{\Delta x^2}. \tag{71}$$

So, owing to Equation (58), Equation (71) reduces to the trivial identity of two reciprocal surfaces $A$ admitting the equality $G\alpha/e = e/m_{Pl}^2 = const$; this equality is reasonable, being mere consequence of the definitions of fine structure constant and Planck mass. Nonetheless, being $e = 4.8\times10^{-10}$ esu and $m_{Pl} = 2.18\times10^{-5}$ g in the cgs system where $G = 6.67\times10^{-8}\ \mathrm{cm}^3\cdot\mathrm{g}^{-1}\cdot\mathrm{s}^{-2}$, one finds

$$\frac{G\alpha}{e} = \frac{e}{m_{Pl}^2} = 1\,\mathrm{cm}^{3/2}/\mathrm{g}^{3/2}\cdot\mathrm{s}. \tag{72}$$

The fact that both ratios are almost exactly equal to 1 is not trivial: in general one would expect simply $G\alpha/e = e/m_{Pl}^2 = const$, with *const* equal to a generic numerical value; the fact that $const = 1$ shows that $G$ and $e$ are linked directly, not via a proportionality constant whose physical meaning should be specifically explained. Equation (72) reveals thus the direct correlation between $e$ and $G$ via $\alpha$, *i.e.* between gravitational and electromagnetic interaction.

On the other hand, despite the simplicity of approach, the compliance of the present model with the relativity, already emphasized by the corollaries of the Section 4, does not appear accidental. This point is elucidated next by four relevant examples.

1) According to Equation (2) $c\Delta t - \Delta x = 0$, whence $c\Delta t - \Delta x' = s$ with the arbitrary length $s \neq 0$ for $\Delta x' \neq \Delta x$. Hence in analogy with Equation (1) $c\Delta t - n'\lambda_C = s$, being $n' = c/v'$. Also, it is possible to write

$$c\Delta t - n'\frac{h}{mc} = s = \frac{mc^2\Delta t^2 - n'h\Delta t}{mc\Delta t}$$

By dimensional reasons, it is also possible to put $n'h\Delta t = m\ell^2$, being $\ell$ an arbitrary length. So, let us show that are definable two invariant equations linked by a square interval of size $\ell_{inv}$: from

$$(c\Delta t)^2 - \ell^2 = sc\Delta t$$

one expects both

$$(c\Delta t)^2 - \ell^2 = \ell_{inv}^2 = \Delta x^2 - \ell^2, \quad sc\Delta t = \ell_{inv}^2,$$

as in effect it is true. Indeed it is possible to express $\ell$ as $\ell = q\Delta x$, with $q$ arbitrary constant; this result reads $(c\Delta t)^2 - \ell^2 = (1-q^2)\Delta x^2$. On the one hand, the right hand sides defines $(c\Delta t)^2 - \ell^2$ in the same reference system of $\Delta x$, *i.e.* it trivially concerns a smaller range. On the other hand, however, $(c\Delta t)^2 - \ell^2$ is not necessarily related to $\Delta x$ via the proportionality constant $1-q^2$ only; being arbitrary by definition, it can be regarded in general as any $\Delta x'$, *i.e.* $\Delta x$ in another reference system mowing with respect to that of $\Delta x$. Hence, the left hand side is an invariant; it also holds therefore for $s\Delta t$. Recalling Equation (9), the time and space invariants with $v = const$ read thus

$$\Delta t' = \Delta t\sqrt{1 - v^2/c^2}, \quad \Delta x' = \frac{\Delta x}{\sqrt{1 - v^2/c^2}}. \tag{73}$$

2) Let $v_1$ be the average velocity component with which one massive particle moves in the range $\Delta x$ and let the particle energy be subjected to fluctuations, during which its velocity component changes to the value $v_2$. Assume therefore that the particle can move in $\Delta x$ at average rates $v_1$ or $v_2$, both arbitrary and allowed to occur during the time range $\Delta x/c$ of Equation (2) because of an unforeseeable fluctuation; the only constrain is that, according to Equation (1), $m \neq 0$ requires $v_1 < c$ and $v_2 < c$. Define therefore

$$\Pi_1 = \frac{1}{n_1} = \frac{v_1}{c}, \quad \Pi_2 = \frac{1}{n_2} = \frac{v_2}{c}, \quad \Pi = \Pi_1 + \Pi_2;$$

the notations emphasize the probabilities $\Pi_1$ and $\Pi_2$ that *m*, delocalized within $\Delta x$, travels just with either velocity before and after the fluctuation. Moreover, $\Pi = v_1/c + v_2/c$ emphasizes that anyway the particle moves, *i.e.* both velocity components are allowed to the particle; this is another way to state that fluctuation in fact occurred. But of course $\Pi$ is effectively definable provided that $v_1/c + v_2/c \le 1$ too, which allows regarding $\Pi$ as pertinent probability that both $\Pi_1$ and $\Pi_2$ are possible for the particle in $\Delta x$. This is the first boundary condition of the present problem. Since this reasoning in $R$ must hold likewise in any other reference system $R'$, it is possible to describe the situation for the range size $\Delta x'$, *i.e.*

$$\Pi = (1+q)\Pi', \quad \Pi = \frac{1}{n_1} + \frac{1}{n_2}, \quad \Pi' = \frac{1}{n_1'} + \frac{1}{n_2'}, \tag{74}$$

where again $q$ is an arbitrary constant. Let the primed and unprimed velocity components be defined in $R'$ and $R$ thinking that in general $\Delta x'$ shifts with arbitrary rate with respect to $\Delta x$. Of course still holds in $R'$ the boundary condition $v_1'/c + v_2'/c \le 1$ in order that the energy fluctuation be regarded in an analogous way in both reference systems. This is the second boundary condition of the problem. Clearly the factor $q > 0$ represents the link between primed and unprimed quantities, *i.e.* it determines the transformation law of velocity components in $R$ and $R'$: in fact the form (74) ensures that if $\Pi \le 1$, then anyway $\Pi' \le 1$ as well. In principle the boundary conditions are unsatisfied simply summing $v_1$ and $v_2$, *i.e.* calculating $\Pi_1 + \Pi_2$; yet the actual form of the sum of velocity components depends on the choice of *q*. Note in this respect that it is reasonable to put

$$q = \Pi_1 \Pi_2;$$

this position ensures that both $v_1$ and $v_2$, whatever they might be before and after the energy fluctuation, are in fact allowed in $R$ and $R'$. In other words, the positions just introduced regard the ratios $v_1/c$ and $v_2/c$ as probabilities of states with and without fluctuation accessible to the particle, and thus in fact occurring, regardless of the choice of reference system. The first Equation (74) reads therefore

$$\Pi' = \frac{\Pi}{1 + \Pi_1 \Pi_2}$$

so that

$$v' = v_1' + v_2' = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}.$$

In this way $v_1 + v_2 \le c$ is always consistent with the probabilistic meaning of $\Pi_1 + \Pi_2$.

This is the well known composition rule of the velocity components along the direction of motion of two reference system reciprocally displacing. The probabilistic meaning of a relevant relativistic property also appears here, already emphasized in demonstrating the perihelion precession [7]. This explains why even the relativistic result is compatible with the quantum approach.

3) It is easy at this point to highlight further the physical meaning of the length $2mG/c^2$. Consider the first Equation (11) and calculate the photon energy $E_0^{ph} = h\nu_0^{ph}$ putting once more $h\nu_0^{ph} = mc^2/n$, in strict analogy with the first Equation (3); so, in particular, $h/2\Delta t = mc^2/n$ too. The left hand side introduces the idea of frequency $1/2\Delta t$ of a photon necessarily confined within a range $\Delta x_{co}$ via the position $\lambda_0^{ph} = 2\Delta x_{co}$; this implies indeed a steady wavelength actually consisting of two half-wavelengths spreading at rate *c* throughout $\Delta x_{co}$. Of course, Equation (3) holds also for the particular photon frequency $\nu_0^{ph}$ related to such wavelength. Replacing once more *h* via the Planck length, $h = 2\pi l_{Pl}^2 c^3/G$ one finds

$$\frac{\omega l_{Pl}^2 c^3}{2G} = \frac{mc^2}{n}, \quad \omega = \frac{2\pi}{\Delta t};$$

then, dividing both sides of $\omega n l_{Pl}^2 c^3 = 2mGc^2$ by $c^4$, the result is

$$\frac{n\omega l_{Pl}^2}{c} = \frac{\omega l_{Pl}^2}{v} = \frac{l_{Pl}^2}{\Delta r} = \frac{2mG}{c^2} = \Delta x_{co}, \quad \omega\Delta r = v.$$

Hence, $2mG/c^2$ of Equation (12) appears again here through the length $\ell = l_{Pl}^2/\Delta r$. The circular frequency $\omega$ shows that the photon cannot escape from the gravitational field of *m*, the photon can only "orbit" around *m* at the black hole distance $\ell$ from the center of gravity. Obviously, this conclusion could be inferred for the second Equation (11) too. The interest to quote here a result already found is that of rising an interesting question: what happens if the photon transits a a distance $\Delta x > \Delta x_{co}$ from *m* greater than that compelling its confinement? The most intuitive answer is that the photon should be simply deviated from its asymptotic straight propagation because of the presence of gravity field of *m*, which reasonably deforms the surrounding space time. Let be therefore according to Equation (12)

$$\frac{2MG}{c^2 \Delta x_{co}} = 1 \tag{75}$$

and write then identically

$$\frac{\Delta x_{co}}{\Delta x} = \frac{2mG}{c^2 \Delta x}. \tag{76}$$

Defining an angle $\phi$ via the arc $\delta s$ of circumference of radius $\Delta x$ such that $\phi = \delta s/\Delta x$, the left hand side becomes $(\Delta x_{co}/\delta s)\phi$. Implement once more a probabilistic approach, noting that $\Delta x_{co} < \Delta x$ is compatible with $\Delta x_{co}/\delta s < 1$. Is in particular interesting the case where $\Delta x_{co}/\delta s = \Pi = 1/2$: indeed $\delta s$ can be defined for a photon coming from minus infinity, approaches m and proceeds towards infinity after being deviated by the gravitational field of m; however this case is physically indistinguishable from that where the photon comes from infinity and proceeds towards minus infinity after an identical deviation. In other words $\delta s$ must be such that $\Pi_{-\infty\to+\infty} = \Pi_{+\infty\to-\infty}$, whereas of course $\Pi_{-\infty\to+\infty} + \Pi_{+\infty\to-\infty} = 1$: *i.e.* the photons is anyway deviated wherever it comes from or proceeds to. Since however either chance only actually happens and has probability $\Pi = 1/2$, the previous result reads

$$\phi = \frac{4mG}{c^2 \Delta x}.$$

This is the well known formula of the light beam bending in a gravitational field, since the angle defining the arc of circumference is equal to that between the tangents to the circumference at the boundaries of the arc, which yield the sought path deviation.

4) Note eventually that $c^2\Delta t^2 - \Delta x^2 = \ell_{inv}^2$ defines an invariant interval $\ell_{inv}$ whatever $\Delta t$ and $\Delta x$ might be. Of course this invariance property of the range size $\ell_{inv}$ holds even if one considers in particular according to Equation (73)

$$c^2\left(\Delta t\sqrt{1-v^2/c^2}\right)^2 - \left(\frac{\Delta x}{\sqrt{1-v^2/c^2}}\right)^2 = \ell_{inv}^2,$$

as in this case both addends at the left hand side remain themselves identically unchanged in two different inertial reference systems in reciprocal constant motion. This expression does not consider the mass of a particle possibly present in the space time. Consider now Equation (75) and introduce an arbitrary distance $\Delta x > \Delta x_{co}$, *i.e.* outside the confinement range, and consider the gravity field of *M* at an arbitrary point outside its event horizon; then Equation (12) yields

$$\frac{2MG}{v_M^2 \Delta x} = 1, \quad \Delta x > \Delta x_{co}, \quad v_M < c,$$

being $v_M$ the local velocity defined by $MG$ just at a distance $\Delta x$. Hence

$$\frac{2MG}{c^2 \Delta x} = \frac{v_M^2}{c^2} < 1. \tag{77}$$

Replacing $v/c$ with $v_M/c$, Equation (77) yields according to Equation (73)

$$c^2\left(\Delta t\sqrt{1-2MG/c^2\Delta x}\right)^2 - \left(\frac{\Delta x}{\sqrt{1-2MG/c^2\Delta x}}\right)^2 = inv_M^2.$$

Simple considerations show that the right hand side reduces to the form $\delta s^2 - \delta r^2 \mathrm{d}\Omega^2$ in spherical coordinates; this is thus nothing else but the metric of the general relativity formulated by Schwarzschild. However this last result shows a crucial difference from the Einstein metrics: the latter assumes that the boundary of the ranges therein appearing are exactly knowable as in the classical physics, the former are instead uncertainty ranges that by definition satisfy the Heisenberg principle [7]. On the one hand this formal analogy explains why the formulas of the general relativity are also found via quantum approach; on the other hand their conceptual difference from the classical physics explains the difficulty of bridging relativistic and quantum ideas. However, the present reasoning shows that the link between quantum and relativistic physics exists indeed and is easily identifiable with the help of elementary considerations. Just this remarkable circumstance allows bridging quantum physics and relativity, despite the Einstein space time metrics is essentially classical physics extraordinarily enriched by the key concepts of four-dimensionality and covariancy of physical laws. In lack of a radically alternative way to infer the relativistic formulas, the mere attempt of modifying or perturbing the standard formulation of the general relativity to bridge deterministic metrics and probabilistic character of the non-real and non-local quantum world, would be difficult or even self-contradicting.

## 8. Conclusions

Starting from elementary considerations, the present model is allowed to describe the fluctuations in a wider theoretical context that includes even relativistic implications. No "ad hoc" hypothesis has been necessary to infer relativistic results, which deserve a few final remarks. The first one emphasizes that in the present context they have been obtained regardless of any preliminary consideration about the covariancy of the physical laws and even about the metrics describing the space time deformation in the presence of matter; actually, instead, the hidden probabilistic meaning of the most famous results of the general relativity is easily acknowledgeable. The second one stresses an open point left by Equation (56) and omitted for brevity taking the absolute value of $F$ in Equation (60), *i.e.* that the space time deformation inherent the time dependence of $\lambda$ could imply in principle contraction or expansion of the range $\Delta x$ and thus both signs of $\dot{\lambda}$; hence, the signs of $F$ correspond not only to attractive or repulsive interaction of the charges $e'$ and $e''$, well known, but also to different chances of gravity force. This point, also remarked in [7] [12], opens a critical problem about the existence of the anti-gravity. This conclusion deserves detailed investigation, too long and complex to be exposed in a short conclusion.

A final remark deserves attention. With little effort and elementary mathematical formalism, Einstein could anticipate himself as done here the most significant discoveries of his general relativity: i.e., as side corollaries of Equation (54) describing the black body fluctuation. Unfortunately his paper [10], despite its great historical relevance, was too purposely focused on the new born Planck physics. May be, the reluctance of Einstein to accept the weird quantum ideas has been the main conceptual obstacle to his opening towards the possible implications of the quantum fluctuations. It is surprising that great intuitions like the photon and the far reaching model of specific heat of solids settled eventually with the mere "hidden variables" of the EPR paradox. The present paper confirms indeed that there is no conceptual gap between quantum and relativistic ideas, as the conceptual foundations of both theories are actually rooted in the quantum concept of space time uncertainty [13].

## References


[1] Siegel, R. and Howell, J.R. (2002) Thermal Radiation Heat Transfer. Volume 1, Taylor and Francis Ed., 7.

[2] Adkins, C.J. (1983) Equilibrium Thermodynamics. Cambridge University Press, Cambridge, 50. http://dx.doi.org/10.1017/CBO9781139167703.005

[3] Carplus, R. and Neuman, M. (1953) *Physical Review Letters*, **83**, 776. http://dx.doi.org/10.1103/PhysRev.83.776

[4] Bergmann, L., Schaefer, C. and Niedrig, H. (1999) Optics of Waves and Particles. W. de Gruyter Ed., Berlin, 595.

[5] Irons, F.E. (2004) *American Journal of Physics*, **72**, 1059. http://dx.doi.org/10.1119/1.1767098

[6] Varro, S. (2006) Einstein's Fluctuation Formula. An Overview. arxiv.org/pdf/quant-ph/0611023

[7] Tosto, S. (2016) *Journal of Advances in Physics*, **6**, 3408-3469.

[8] Amelino-Camelia, G. (2002) *Physics Letters B*, **528**, 181-187.

[9] Dandliker, R. (2000) The Concept of Modes in Optics and Photonics. 6th *International Conference on Education and Training in Optics and Photonics*, **3831**, 193-198.

[10] Einstein, A. (1909) *Physikalische Zeitschrift*, **10**, 185-193.

[11] Landau, L. and Lifshits, E. (1966) Theorie du Champ (in French). Ed MIR, Moscow.

[12] Tosto, S. (2013) *Physics International*, **4**, 135-151. http://dx.doi.org/10.3844/pisp.2013.135.151

[13] Tosto, S. (2014) *International Journal of Physics and Astronomy*, **27**, 1136.

Journal of Modern Physics, 2016, 7, 1702-1711
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Why Does Newton's Apple Fall *Vertically* to the Ground: The Gravitation Code

**Jean-Paul Auffray**

Ex: Courant Institute of Mathematical Sciences, New York University, New York, NY, USA
Email: jpauffray@yahoo.fr

**How to cite this paper:** Auffray, J.-P. (2016) Why Does Newton's Apple Fall *Vertically* to the Ground: The Gravitation Code. *Journal of Modern Physics*, **7**, 1702-1711.
http://dx.doi.org/10.4236/jmp.2016.713153

**Received:** August 23, 2016
**Accepted:** September 20, 2016
**Published:** September 23, 2016



## Abstract

No, Isaac Newton did not "explain" gravitation. What he did, and this certainly constituted all and by itself a great achievement, was to recognize (to "assert") the universal character of gravitation: all material objects (bodies) attract each other by gravitation. But how does gravitation perform its deeds? This remained a mystery to Newton. In a "desperate move" at the end of his life, he introduced the concept of "Particles which are moved by certain *active Principles* [our emphasis]—such as is that of Gravity" he said. We resurrect this scheme, we provide it with a quantum structure—a stunning new insight into the workings of gravitation obtains.

## Keywords

Gravitation, Newton, Fatio de Duillier, Ethereal Particles, I-Meteons

## 1. Introduction

Ever since in 1896, the great Swedish physical chemist Svante August Arrhenius (1859-1927) propagated the fallacious assertion making carbon dioxide ($CO_2$) a dangerous "Greenhouse Effect" gas contributing to climate warming [1]. Physicists at large have displayed a startling ability to believe in just about anything. Astute observer of human nature, Albert Einstein, took advantage of this to cause physicists surreptitiously to see in him the greatest physicist who ever lived—after Newton, of course. Speaking of Newton, who has not heard of the legendary story of the future great Natural philosopher watching an apple fall off a tree to the ground under the influence of that mysterious force, Gravitation, deciding then that he would be the one, some day, who would unravel the secrets of gravitation—decode the Gravitation Code—, a natural ambition for someone who, like himself, was destined to become a devoted theologian and alchemist [2].

As it turned out, Newton did *not* break (decode) the Gravitation Code. What he did, and this certainly constituted all and by itself a great achievement, was to recognize (to "assert") the universal character of gravitation: all objects (bodies) having mass attract each other by gravitation. If one of the two bodies is the Earth, gravitation is conventionally called *gravity*. This constitutes, by the way; a circular reasoning: any two bodies which have mass attract each other by gravitation; any two bodies which attract each other by gravitation have mass.

Very well, but how does gravity (gravitation) operate? This remained a frustrating puzzle to Newton to the end of his life. In an ultimate "desperate move"—and this might have been in the wake of his encounter a few years earlier with a young genius mathematician who *did* propose a "Mechanical" explanation of the workings of gravitation—Newton introduced the concept of "*Particles* which are moved by certain *active Principles*—such as is *that of Gravity*", he said (our emphasizes) [3].

Gravity caused by "*active principles*"… Newton meant an action arising from the "Will and the Spirit" of God, *i.e.* not caused by any mechanical, material agent.

We retrieve this commonly ignored Newtonian scheme. We provide it with a quantum formulation. A breathtaking revolutionary insight into the (hidden) workings of gravitation obtains.

## 2. Birth of a Mechanical Model of Gravitation

### 2.1. Birth of a Genius

In 1683, after twenty years sharing the same rooms in Trinity College at the University of Cambridge, Newton's companion John Wickens left Cambridge to become a country vicar somewhere else. Henceforth on his own to pursue in secrecy his alchemical experiments and his arduous attempts to decipher the prophecies consigned in the Old Testament, Newton undertook to write the first Book of the philosophical treatise which was to make him famous, the renowned *Principia*. In the midst of these intense activities, he suddenly became involved with a young genius mathematician twenty-two years younger than him.

Born the seventh child in a family of fourteen siblings originally established in the small town of Chiavenna in northern Italy, the boy is eight-years-old in 1672 when his father, Jean Baptiste Fatio, or Faccio short for Bonifaccio, acquires a new residence in the canton de Vaud in Switzerland, the Seigneurie de Duillier (**Figure 1**), whereby his son Nicolas assumed proudly henceforth the name Nicolas Fatio *de Duillier*.

Talented, ambitious and adventurous, young Nicolas seeks from the start the company of men of high intellectual standing, preferring them to "little Persons". At nineteen, he is in Paris working under the leadership of the great astronomer Domenico Cassini at the Observatoire royal, but the 22 October 1685 the King of France, Louis XIV, who used to sign his name *Nous Louis Roi* (We Louis King) promulgates the Edict of Fontainebleau which forbids the practice of Protestantism on the territory of his kingdom. Young Nicolas takes refuge in the Netherlands in the company of the "excellent Mathematician and good Philosopher" Christiaan Huygens, who soon takes

**Figure 1.** Le château de Duillier.
https://upload.wikimedia.org/wikipedia/commons/thumb/a/a4/Chateau_Duillier2.jpg/420px-Chateau_Duillier2.jpg

him with him to London, where Fatio is promptly elected a Fellow of the *Royal Society*. He is twenty-four years old. Newton himself had been elected a Fellow of the Royal Society sixteen years earlier.

The encounter between the two Fellows left them both in a daze: young Fatio informed his elder that he had elaborated a "Mechanical" theory of gravitation!

Written in seventeenth century French, Fatio's *Mémoire Sur la Cause de la Pesanteur* [4]—Essay on the Cause of Gravity—remained unpublished during Fatio's lifetime. It was retrieved among his papers after his death at his residence in Maddersfield near Worcester in England on May 10$^{th}$ 1753.

## 2.2. Fundamental Assumptions

The starting points of Nicolas Fatio's mechanical theory of gravitation are four. We first list them as formulated by Nicolas Fatio in his original language, then express them using modern words and concepts. They are:

1) The World—the Universe, the Cosmos—is composed of (apparently) solid bodies—Fatio calls them "coarse bodies"—which in realty are composed of "Atoms" which are "porous"—*i.e.* full of (invisible) holes.

2) Atoms have geometrical shapes which make them look the same in all directions—they are *anisotropic.*

3) Besides "coarse bodies", the World—the Universe, the Cosmos—also contains another kind of matter—Fatio calls it *ethereal.*

4) *In fine*, the coarse bodies present in the Universe are so porous that the ethereal particles can move freely through them, generating in the process the phenomenon we observe as gravitation.

These assumptions take on added significance when expressed in contemporary terms. Astrophysicists tell us indeed nowadays that the observable universe contains a phenomenal number of *galaxies*, each composed of billions of suns (*stars*), themselves agglomerates of atoms themselves made up of so-called *elementary particles.* If the observable universe—if the Cosmos—is made up of elementary particles—electrons, quarks, neutrinos and the rest—which individually occupy essentially no space, then, *in fine*, against all appearances said universe is fundamentally a huge vacuum—a *Void*— dotted of point-like "nothings", a thought that Fatio expressed in his days in these words: "*Je suppose que les differents Espaces du Monde sont presque entierement Vuides* [sic] *de Matiere*. —I assume that the different Spaces in the World are almost entirely Void of Matter."

Let us investigate these considerations in the context of a specfic example.

### 2.3. The Special Case of Newton's Falling Apple

Consider the legendary example of Isaac Newton as a boy observing an apple fall off a tree down to the ground (**Figure 2**). To us, ordinary inhabitants of the planet Earth, nothing special about this: grass is green, the sky is blue, and ripe fruits fall off branches. Newton was more subtle: to him the Earth as a whole does not "attract" the apple; instead, the apple "sees" the Earth's *Center*, "which is a mathematical point", he said; the apple is attracted by it, or to it, and moves toward it—an Act of God, by Newton's reckoning.

This is where Nicolas Fatio's genius intervenes. Let us idealize the apple by representing it as a small "coarse" solid sphere. By the criteria exposed in our Heading 2.2., this sphere is really a ghost shell containing porous Atoms immersed in the "Vuoid", as Fatio spells it. Essentially empty, this ghostly sphere is nevertheless capable of experiencing the effects of gravitation. How does it do that?

Young Fatio imagined the following clever scheme. In addition to "coarse matter", he said, the world—the universe, the cosmos—also contains other kinds or species of matter, and he described their qualities…

**Figure 2.** Young Newton observing the falling apple.
http://www.akg-images.co.uk/Docs/AKG/Media/TR5/0/3/3/f/AKG333767.jpg

Fatio wrote his memoir using a vocabulary composed of seventeenth century French words which are difficult to "translate" properly in any contemporary idiom. Giving up word-for-word translation, we express in modern terms Fatio's fundamental concept of the existence in the universe of at least two distinct kinds of matter—*coarse* matter and *ethereal* matter according to one of his choices of qualifiers.

I-meteons—*i-points* in motion—constitute ideal quantum-world candidates to incarnate the "ethereal" or "second-nature" particles said by young Fatio to be responsible for gravitational effects. We invite the reader unfamiliar with the i-point and i-meteon concepts to consult the presentation of their significances in [5].

As per this contemporary quantum scheme, i-meteons are generated randomly, *i.e.* they propagate equally (indifferently) in all directions in Absolute space, the only restriction to their activity being that they each carry precisely one element—one unit, one *quantum*—of "motion", a designation introduced to replace the term "dynamical action" ill-appreciated by theoreticians in general nowadays (they prefer to speak of "energy").

Let our idealized spherical "coarse apple" be on its own, isolated somewhere in the cosmos far from the influence of any other "coarse body" such as the Earth. Our apple is then exposed "naked" so to speak to the random flux of i-meteons arriving from all parts of the world (**Figure 3(a)**). Arriving randomly from all directions, the i-meteons exert no net "push" on the apple in any one particular direction—no gravitational pull or push is generated on our "apple".

Now allow another "coarse body" to be in the viccinity of the apple, say next to it for the purpose of illustration. By Nicolas Fatio's contention—and this is the key to the proper functioning of his Mechanical model of gravitation—, being next to each other the two coarse bodies "filter" the ethereal Particles—the i-meteons for us—arriving on them from one side (**Figure 3(b)**). An asymetry or imbalance is generated in the system, each of the two bodies is impiged less on one of its sides and starts moving in that direction, being apparently gravitationally "attracted" by the other body whose

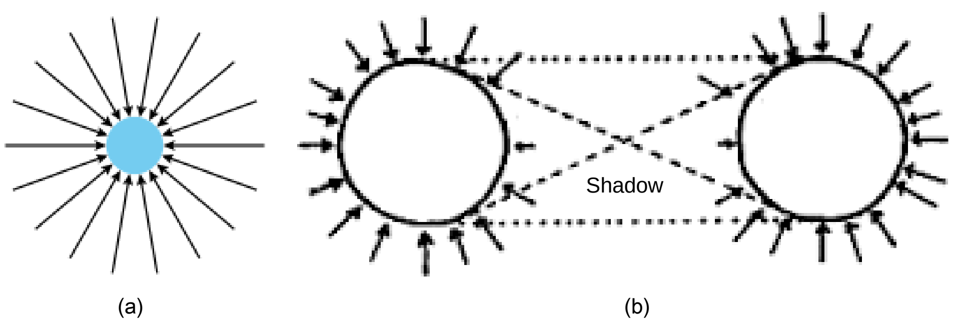



**Figure 3.** (a) I-meteons impinging randomly from all sides on an isolated coarse spherical body generate no net "push" on it. https://upload.wikimedia.org/wikipedia/commons/thumb/2/2d/Pushing1.svg/225px-Pushing1.svg.png (b) Coarse bodies "shadowing" each other. https://encrypted-tbn1.gstatic.com/images?q=tbn:ANd9GcRK4tfCKk7igUhiCcKXgXQJaZv7L6Y76rq6kWJwvInTWBGpgf2C

presence causes the imbalance to occur. In reality each of the two bodies is driven toward the other by the "push" of unbalanced incoming i-meteons—etheral Particles for Nicolas Fatio.

And suddenly, contemplating Figure 3(b), a simple, startling "explanation" as to why Newton's apple falls v*ertically* to the ground and why we are held *vertically* when we stand erect on the ground somewhere—anywhere—on the surface of the Earth minding our own business emerges. Unseen but nevertheless present around us, i-meteons impige on our "coarse body" just about equally from all directions… from all directions but one: the massive coarse Earth under our feet prevents i-meteons from entering our body vertically upward through our feet. As a result the i-meteons which enter our coarse body vertically downward through the top of our head are not compensated by i-meteons entering our body from the opposite direction; they exert a net push vertically downward on our body—we stand *vertically* erect.

It remains to account for the quantitative aspects of the phenomenon.

## 3. Probing the Model

### 3.1. Why the Inverse Square Law

Gravitation is known to possess two fundamental characteristics: 1) it obeys an "inverse-square" law and 2) it satisfies the law of "mass proportionality". Does our quantum "shadow model" of gravitation satisfy these two requirements?

The imbalance in i-meteon distribution is independent of the size of the enclosing sphere while the sphere surface area increases as the square of the radius. The imbalance per unit area decreases inversely as the square of the distance between the two body centers: the inverse-square law obtains.

### 3.2. Why Mass Proportionality

Nicolas Fatio spent three years of his life thinking about it, he said… until he came up with a plausible answer while in London in the Fall of 1689. The key of his discovery is this.

While crossing a "coarse body", an ethereal particle can experience two kinds of mishaps: 1) be "absorbed"—Fatio calls it be "condensed"; or 2) loose some of its momentum—Fatio calls it loose some of its "Mouvement" (Motion). "Je reconnus que cette Condensation, he said, etoit aussi petite qu'on vouloit, de même que la Perte du Mouvement, jusques à devenir infiniment petite, si l'on faisoit les Suppositions necessaires."—I found that the Condensation [Absorption] was as small as one might want, as well as the loss of momentum, so as to become infinitesimally small, if one made appropriate Suppositions [Hypothesises].

Anticipating the modern belief in the existence in nature of elementary particles, he postulated the presence in coarse bodies of several "Orders" of particles which attract each other by gravitation, except for the particles belonging to one particular Order. The particles belonging in this Order are so perfectly 'hard", he said, that their "Ressort" [ability to reflect off an impinging entity] can be thought to be infinite, the Cause of which is "Metaphysical in essence" and has its origin in "the Will of the Creator", a belief he shared with Newton [2].

The fundamental question which arises in the frame of this scheme concerns the nature of the interaction which takes place when an ethereal Particle—an *i-meteon* for us—goes through the "Vuoid" contained in a given coarse body. Nicolas Fatio's key assumption in this regard stands in one sentence (our translation): "Gravity is produced, in my view, by Exceedance of the Speed of the Particles of this [ethereal] Matter which impinges on the Earth for example, or some coarse Atom of which it is composed, over their Speed when they are reflected…" [4]. In brief: incident ethereal Particles strike the body at a very high speed, then rebound with a slightly lower speed.

Advantages; This is actually a more sophisticated model than simply assuming total absorption because it reognizes that perfect reflection would result in no anisotropy at all in the surrounding flux, and therefore no net force of gravity. Instead it allows for a combination of reflection and absorption of momentum and avoids mass accumulation (Fatio assumed his ethereal Particles to have mass). He also stressed the fact that to produce a given amount of gravity, we can suppose the bombarding Particles to be arbitrarily small while assuming their speeds to be arbitrarily great. This automatically diminishes the drag induced by the movement of coarse bodies to a negligible amount. Fatio also argued that by supposing the speed of the ethereal Particles to be extremely great, the amount by which they are slowed can be made as small as we wish, so there need be no appreciable dimunition of their agitation over time.

These considerations fare well with our quantum scheme as described in [5]. When the ubiquitous quantum sets in motion one of the points the dimensionless Void contains, this point becomes a e-meteon if the quantum sets it to move at the speed of light c; a i-meteon if the quantum sets it to move at speeds slower or faster than the speed of light—in brief at speeds not related to the speed of light.

This concept will come as a shock to those who believe the speed of light to be the ultimate speed that a body can achieve while moving in the universe—a material body, yes, but i-meteons are not material bodies, they are i-meteons, *i.e.* they are points.

In this regard, we call attention to the invention made by the incomparable Richard Feynman, Nobel Laureate for Physics in 1965, when he postulated in 1969 the existence in nature of point-like *partons* defined with respect to a physical scale making it possible for them to exist either as "valence constituents" of elementary particles, or as "nonvalence partons" forming a "sea" [6].

Two kinds of point-like partons…. We will not be surprised to find in this magical invention matter useful for enriching our own invention of i-meteons and help in understanding the nature of their behavior in interactions. We shall leave these considerations for further studies to be conducted.

## 4. Sad Ending

After French King Louis XIV revoked the Edict of Nantes in 1685, a group of radical Protestants, the *Camisards,* began a violent insurrection in the French countryside. In

1706, after the movement was put down in France, some of the Camisards immigrated to England where they became known as the *French Prophets*. Fatio became a disciple of their leader, Elie Marion. He went with him into animated visionary trances during public sermons, claiming to perform miracles (including raising the dead), and made extravagant prophecies of the imminent end of the world. In 1707 Marion, Fatio, and another French Prophet were convicted of blasphemy and sedition and sentenced to be pilloried for two days. A sign placed over Fatio's head explained the reason(s) for his being exposed to the pillory (**Figure 4**).

"Nicolas Fatio convicted for abbeting and favouring Elias Marion, in the Wicked and counterfeit prophecies, and causing them to be printed and published, to terrify the Queen's people."

In 1710 the French Prophets left England for Holland, where Nicolas Fatio was twice more sentenced to the pillory. He accompanied Marion thereafter on travels through various European countries, attempting to make converts. While in Turkey in 1712, Marion fell ill, and died. Fatio returned to England, settling near the town of Worcester where he remained for the rest of his life, meditating on the prophecies and pursuing his scientific research. After his death on May 12, 1753 (he was then 89-years old), Swiss mathematician Georges-Louis Le Sage, himself a Huguenot, visited Fatio's former English estate in Maddersfield and retrieved his gravitation papers. In 1784 he made Fatio's gravitation theory known under his own name, asserting that the ethereal particles responsible for gravitation actually originated in another world, deserving to be called accordingly "ultramundane corpuscules". But this is another story, outside the scope of

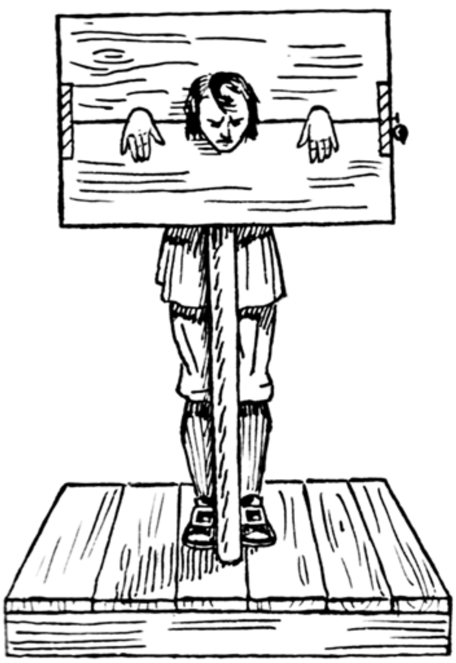


**Figure 4.** Pillory.
http://media.virbcdn.com/cdn_images/resize_1024x1365/0f/ContentImage-9524-287624-545pxPillory_PSF.png

the present note.

## 5. Conclusions: Breaking the Gravitation Code

Newton recognized ("asserted") in his days the universal character of gravitation, but he never found nor proposed an explanation for it. He ended up believing gravitation to be the result of a (permanent) "Act of God" and he was sharply critical of those who thought otherwise. "That gravity should be innate, inherent, and essential to matter, he wrote, so that one body may act upon another at a distance through a vacuum, without the mediation of anything else, by and through which their action and force may be conveyed from one to another, is to me **so great an absurdity** [our emphasis], that I believe no man who has in philosophical matters a competent faculty of thinking, can ever fall into it." To which he added this verdict: "Gravity must be caused by an agent acting constantly according to certain laws; but whether this agent be material or immaterial, I have left to the consideration of my readers." [3]

An "agent acting constantly according to certain laws"… *the quantum* to us in this note.

Great mathematicians—among them James Maxwell and French mining engineer genius Henri Poincaré, inventor in 1900 of the famous relation $E = mc^2$—have shown convincingly that classical collisions of incoming particles of some sort with the particles constitutive of matter could not legitimately account for gravitation [7].

I-meteons are quantum *points*, not "particles". Their interventions in the affairs of the System of the World—at speeds not related to the "speed of light"—cannot be assimilated to those of the plain particles described in the Standard Model. Indeed, a new era in our understanding of the way motion (action) models the cosmos has been initiated and awaits further investigations [8].

In his dreams, Newton decided "Gravity must be caused by an agent acting constantly according to certain laws". To us in this note, the agent Newton envisioned is the ubiquitous quantum. We shall leave it happily at that and salute with respect the memory of the brilliant young genius ex-Camisard sympathizer, ex-*French Prophet* who showed the way to Newton… and to us. Wouldn't it be a wonder if, thanks to him, the Gravitation Code had at last been broken?

## Acknowledgements

We wish to express our gratitude to Joline (Mésange) Boekschoten, and to Ms. Han Xu (Hellen), JMP Editorial Board Assistant, for her valuable help and kind advice in properly preparing this note for publication in this journal.

## References

[1] Arrhenius, S. (1896) *Philosophical Magazine and Journal of Science Series* 5, **41**, 237-276. http://www.rsc.org/images/Arrhenius1896_tcm18-173546.pdf

[2] https://en.wikipedia.org/wiki/Isaac_Newton%27s_occult_studies

[3] Newton, I. (2007) Original Letter from Isaac Newton to Richard Bentley.

http://www.newtonproject.sussex.ac.uk/view/texts/normalized/THEM00258

[4] Fatio de Duillier, N. Sur la Cause de la Pesanteur. http://www.mahag.com/grav/bopp/fatio-bopp.pdf (in French). http://www.mathpages.com/home/kmath041/kmath041.htm (In English).

[5] Auffray, J.-P. (2015) *Journal of Modern Physics*, **6**, 878-889. http://dx.doi.org/10.4236/jmp.2015.67092

[6] Feynman, R.P. (1969) The Behavior of Hadron Collisions at Extreme Energies. High Energy Collisions: Third International Conference at Stony Brook, N.Y. Gordon & Breach, 237-249.

[7] http://www.mathpages.com/home/kmath209/kmath209.htm

[8] Auffray, J.-P. and El Naschie, M.S. (2016) *Journal of Modern Physics*, **7**, 156-161. http://dx.doi.org/10.4236/jmp.2016.71017

Journal of Modern Physics, 2016, 7, 1712-1724
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Characteristics of AlGaN/GaN HEMTs for Detection of MIG

**Hasina F. Huq, Hector Trevino II, Jorge Castillo**

Department of Electrical Engineering, University of Texas-Rio Grande Valley, Edinburg, TX, USA
Email: hasina.huq@utrgv.edu

**How to cite this paper:** Huq, H.F., Trevino II, H. and Castillo, J. (2016) Characteristics of AlGaN/GaN HEMTs for Detection of MIG. *Journal of Modern Physics*, **7**, 1712-1724.
http://dx.doi.org/10.4236/jmp.2016.713154

**Received:** August 19, 2016
**Accepted:** September 20, 2016
**Published:** September 23, 2016



## Abstract

The aim of this research work is to analyze the surface characteristics of an improved AlGaN/GaN HEMT biosensor. The investigation leads to analyze the transistor performance to detect human MIG with the help of an analytical model and measured data. The surface engineering includes the effects of repeatability, influence of the substrate, threshold shifting, and floating gate configuration. A numerical method is developed using the charge-control model and the results are used to observe the changes in the device channel at the quantum level. A Self-Assembled Monolayer (SAM) is formed at the gate electrode to allow immobilization and reliable cross-linking between the surface of the gate electrode and the antibody. The amperometric detection is realized solely by varying surface charges induced by the biomolecule through capacitive coupling. The equivalent DC bias is $6.99436 \times 10^{-20}$ V which is represented by the total number of charges in the MIG sample. The steady state current of the clean device is 66.89 mA. The effect of creation and immobilization of the protein on the SAM layer increases the current by 80 - 150 μA which ensures that successful induction of electrons is exhibited.

## Keywords

AlGaN/GaN, HEMT, WBG, Biosensor, MIG, Charge-Control Model

## 1. Introduction

Over the past several years, there has been much research into developing newer and less invasive ways to monitor and detect several different biological cells and molecules [1]-[6]. Such biological elements include but not limited to proteins, enzymes, antibodies, and tissue cells. The need for such development arises from the current method [7] [8]. With advances in medicine and technology, a growth in understanding of key biomolecules play certain roles and functions in the development of the diseases and it is becoming more useful in the development of such electronics devices [5]-[8]. AlGaN/GaN based HEMT devices have become very attractive in the world of biological modified field effect transistors (BioFETS/biosensing) due to their thermal stability, high-sensitivity, and label-free/real time detection. They also exhibit chemical inertness to extreme sensing environments [1]-[6]. The unique ability of GaN material is to exhibit spontaneous and piezoelectric polarization (~1200 - 1500 $cm^2$/V-S) in heterojunctions without any need for material doping [8]-[10].

Kang *et al.* reported that the close proximity of this layer to the surface (<35 nm) is extremely sensitive to the ambient changes in surface charge and it results in greater detection sensitivity [2] [4] [11]. There exists a substantial amount of work on analytical and empirical modeling of the devices. However, a few of these models address the issue of the characteristics of the SAM layers. The effects of repeatability influence of the substrates, and threshold voltage shifting are also the key parameters in designing such bio sensor.

Monokine induced by interferon gamma (CXCL9/MIG) is a critical biological marker for determination of transplant rejection [12]-[21]. The range of concentration in normal disease states is approximately 0.2 - 3 ng/mL (or 40 - 100 pM) while concentration in pathophysiological disease states is 10 - 400 ng/mL (or as high as 34 nM) [22] [23]. It is a highly charged particle, having net 20 positive charges per molecule [23] [24]. Early detection of this key biomarker is significant and can result in quicker/appropriate treatment. Preparation of such devices is rigorous, and due to the fragile nature and small scale of the HEMT and minute quantities of analytic solutions, careful preparation must be exercised to create a customized biosensor.

## 2. Methodology

An accurate and robust analytical and empirical model is imperative for predicting the device performance in thiol chemistry. In order to detect MIG, the gate electrode of the HEMT must be functionalized using thiol chemistry. Utilizing the gold-plated surface electrodes of the device, a self-assembled monolayer (SAM) is developed. This SAM layer consists of a crosslinker, (dithiobis succinimidyl propionate (DSP)) which forms a strong chemisorption bond with the gold surface. The linkage formed between DSP and the gold surface is very stable, exceeding the strength and stability of covalent silane bonds with glass [25]. An antibody for the target analyte (Anti-MIG) is then introduced to the gate surface and immobilized by the other end of the crosslinker. Anti-MIG is a negatively charged molecule and upon binding with DSP, an increase in drain current occurs, as the positive surface charge potential is altered and the resulting sheet carrier concentration in the hetero-interface is influenced [26] (**Figure 1**). After this step, the device is ready to use. The Anti-MIG will only interact with MIG and upon introduction of the analyte will bind to the immobilized Anti-MIG. The positively charged MIG pairs with the negatively charged Anti-MIG, neutralizes it, and the resulting activity alters the conductivity of the channel by changing the charge distribution in the conjugated molecules (**Figure 1**).

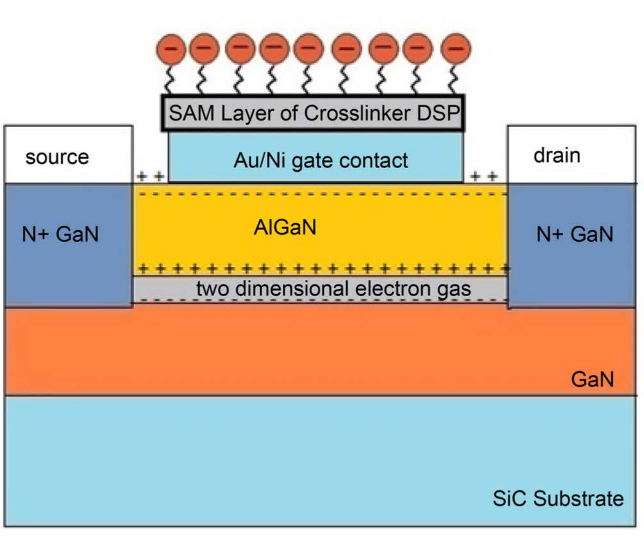



**Figure 1.** Visualization of chemically prepared device before and after conjugation has taken place. Charges due to spontaneous and piezoelectric polarizations are displayed to visually represent how surface charges are impacted by the chemical preparation of the device.

This produces an observable decrease in drain current. The charges induced by these events are by way of capacitive coupling and therefore are analogous to the application of a DC bias at the gate surface [4] [27]-[31]. The SAM layer consists of the crosslinker (DSP) and the immobilized antibody (Anti-MIG). Charges due to spontaneous and piezoelectric polarizations are displayed to visually represent how surface charges are impacted by the chemical preparation of the device.

## 3. Model Formulation

MIG is positively correlated with transplant rejection and has been shown to have about net 20 charges per molecule at a pH concentration of 7.4 (the normal concentration of human blood) [23] [24]. Assuming a disease concentration of 34 nM, the elementary charge is $1.6 \times 10^{-19}$ C. The number of charges per molecule and the total number of molecules in diseased states are calculated by using Avogadro's number. The step by step process can be seen in the following series of equations.

$$\left(34\times10^{-9}\,\text{mol}\right)\times\left(6.022\times10^{23}\,\text{mol}^{-1}\right)=2.04748\times10^{16}\,\text{molecules}\quad [23]\ [24] \tag{1}$$

$$2.04748\times10^{16}\,\text{molecules}\times20\ \text{charges per molecule}$$
$$=4.09496\times10^{17}\ \text{total charges present in diseased state}$$

Recombinant MIG is a complex protein consisting of about 103 amino acid residues with a predicted molecular mass of 11.7 - 12 kDa [32]-[34]. Using solely the molecular mass, the number of charges per vial of sample solution used in experimentation (each vial contains 5 μg/mL of sample) can be determined by first finding the molarity of the solution. Then, the number of molecules in the sample size is found and associated a charge with each molecule independently. The following equations are used to determine the total number of charges per vial: $\left(0.000005\ \text{g}\right)/\left(11700\ \text{g/mol}\right)=427.35\ \rho\text{mol}$,

$$\left(427.35\ \rho\text{mol}\right)/\left(0.001\ \text{L}\right)=4.27.35\ \eta\text{M in a 5 }\mu\text{g/mL sample of solution} \quad (2)$$

Then, solving again for the number of charges per mole it is obtained:

$$\left(427.35\times10^{-9}\,\text{mol}\right)\times\left(6.022\times10^{23}\,\text{mol}^{-1}\right)=2.5735\text{ molecules}$$

$$2.5735\text{ molecules}\times20\text{ charges}=5.147\times10^{18}\text{ total charges} \quad (3)$$

Given that 0.3 mL samples were used during experiment:

$$5.147\times10^{18}\text{ charges}\times\left(1/3\right)=1.71567\times10^{18}\text{ total charges per experiment}$$

Multiplying these total charges by + 1e, we obtain:

$$\left(1.71567\times10^{18}\right)\times\left(1.6\times10^{-19}\,\text{C}\right)=0.274507\text{ C of charge total}.$$

In solid state physics, electron-volts (eV) are used to represent a unit of kinetic energy obtained by accelerating an isolated electron across a potential difference of 1 volt. Thus, 1 eV is equivalent to 1 electric charge times one (1 * e) Joules. Since 1 volt = 1 Joule/Coulomb, then 1 Coulomb = 1 Joule/volt. Since the proteins are immobilized on the floating gate surface, they directly influence surface charges which have an impact on interface charges. According to the work published by Abou-El-Ela (2013), the electron transport characteristics in Wurtzite blend GaN are compared and contrasted amongst varied Electric fields. It is demonstrated that electron velocity in GaN, in the absence (or very small value) of an applied Electric Field possesses velocities in the range of $0.25 \times 10^5$ m/s and $1.25 \times 10^5$ m/s and exhibit an allowable range of energies of ~0.12 eV [35] [36]. Therefore, the DC bias can be calculated as:

$$\text{Charge}=\text{Energy/Voltage}\rightarrow0.274507\text{ C}=\left[\left(1.6\times10^{-19}\right)\times\left(0.12\right)\text{ Joules}\right]\Big/\text{V} \quad (4)$$

Solving for voltage yields V = $6.99436 \times 10^{-20}$. This is the equivalent DC bias represented by the total number of charges in the MIG sample.

## 4. Numerical Model

AlGaN/GaN HEMTs on SiC, Sapphire and diamond substrates are designed using ATLAS™, DECKBUILD™, and TONYPLOT™ by SILVACO™. Charge-Control Analytical models for AlGaN/GaN HEMTs are used to develop the HEMT devices. The HEMT device consists of a 200 Å undoped AlGaN layer (with a concentration of 0.18) which is grown on a 1.5 µm GaN layer. A 400 Å buffer layer consisting of 150 Å AlN layer, and a 250 Å AlGaN layer is developed on a 2 µm thick substrate. A gate length of 2 µm is used. A visualization of this device can be seen in Figure 2(a) along with the corresponding band diagram taken at the heterojunction Figure 2(b). The changes in the device channel at the quantum level are observed. Real-time amperometric responses are also observed with the help of a designed circuit.

The threshold voltage is developed using the Charge-Control Analytical models, Albrecht's equations are used to model the knee voltage. Figure 3 compares the current-voltage behavior of the HEMT devices. The threshold voltages from the simulation results are in close agreement to the threshold voltage from the experimental results.

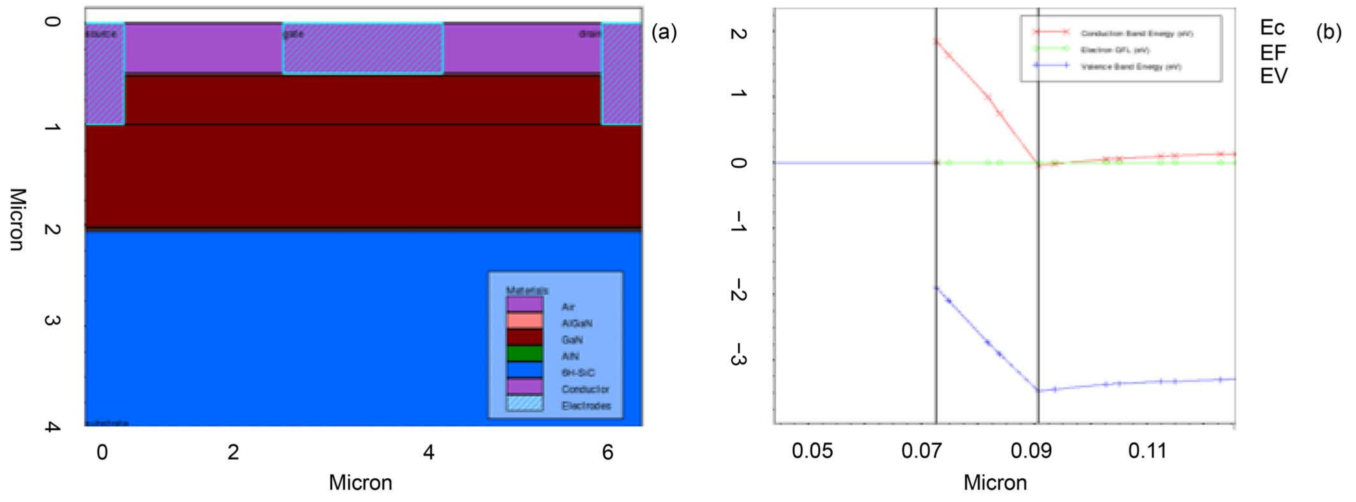



Figure 2. (a) Modeled device using SILVACO™, (b) band diagram of HEMT.

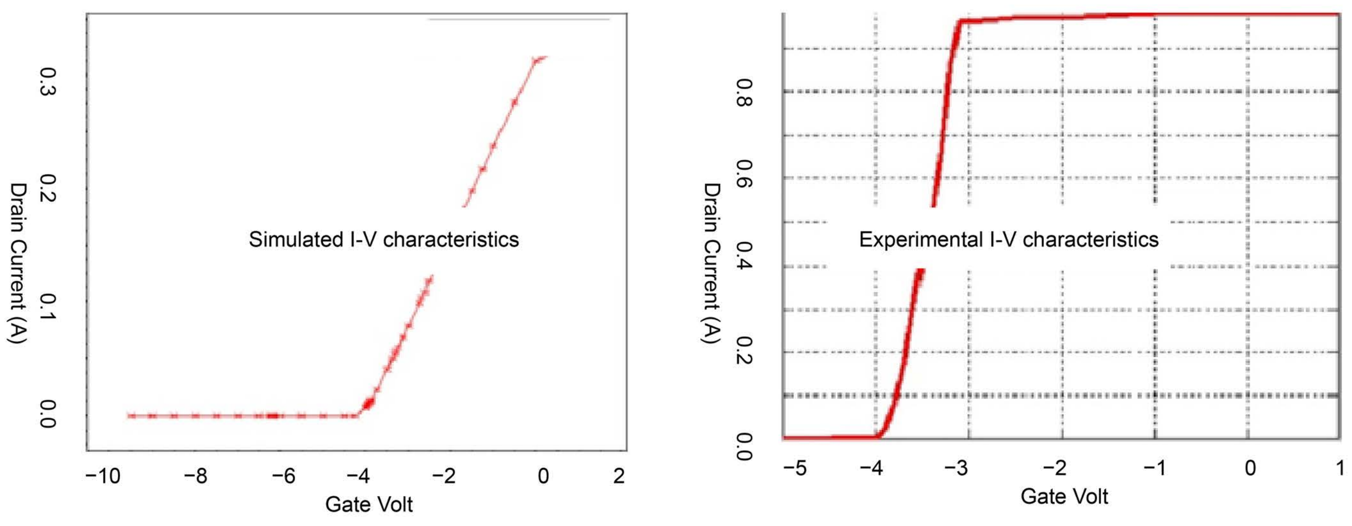



Figure 3. Simulated threshold voltage vs actual device threshold voltage.

However, the saturation drain current is higher in experimental results. The reason is that even though the analytical model may not be accurate from the point of view of absolute value of the measured curve, it gives a rough estimation of how the electrical characteristics change when the parameters are modified.

The threshold voltage is developed using the Charge-Control Analytical models, Albrecht's equations are used to model the knee voltage. Figure 3 compares the current-voltage behavior of the HEMT devices.

The threshold voltages from the simulation results are in close agreement to the threshold voltage from the experimental results. However, the saturation drain current is higher in experimental results. The reason is that even though the analytical model may not be accurate from the point of view of absolute value of the measured curve, it gives a rough estimation of how the electrical characteristics change when the parameters are modified.

The Output characteristics of the simulated AlGaN/GaN HEMTs with different $V_{GS}$ = −3, −2, −1, −0, 1. V are shown in Figure 4. The derived DC voltage is applied to a floating gate configuration on SILVACO™ to simulate the effect of creation and immobilization of the protein on the SAM layer and an 80 - 150 µA increase in current is observed on various substrates with VGS-4 to 1 V and VDS 0 - 4 V. Figure 5 shows the results where the drain current increases about 150 µA. Since Anti-MIG and MIG are equal and opposite in charge, an additional DC bias is modeled with a charge approximately equal to 86.7% of the initial charge bias.

Additional observations are seen at the heterojunction interface before and after conjugation. As expected, a change in current density at the interface is observed (Figure 6).

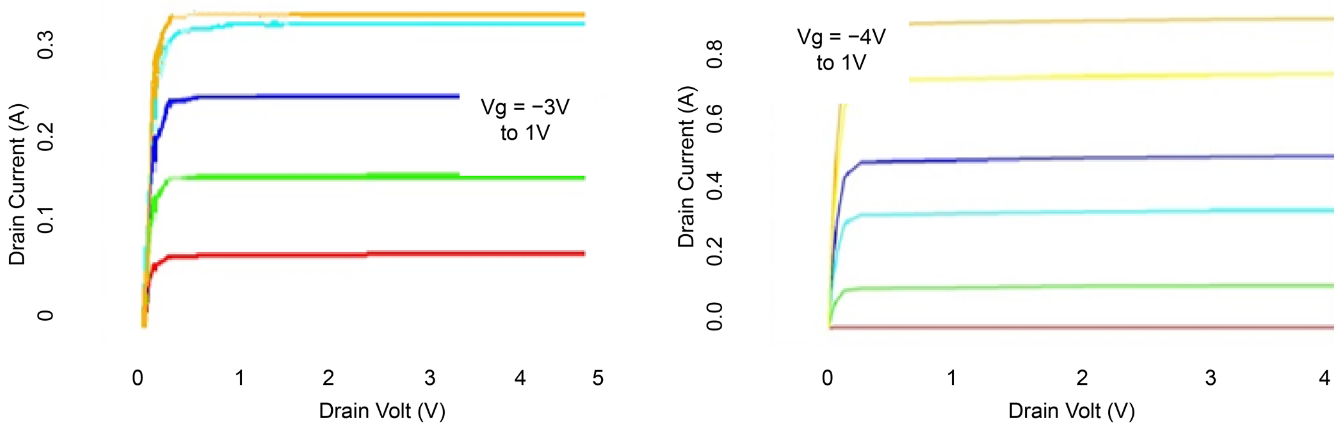



Figure 4. Output characteristics of simulated AlGaN/GaN HEMTs with different $V_{GS}$ = −3, −2, −1, 0, 1 V.

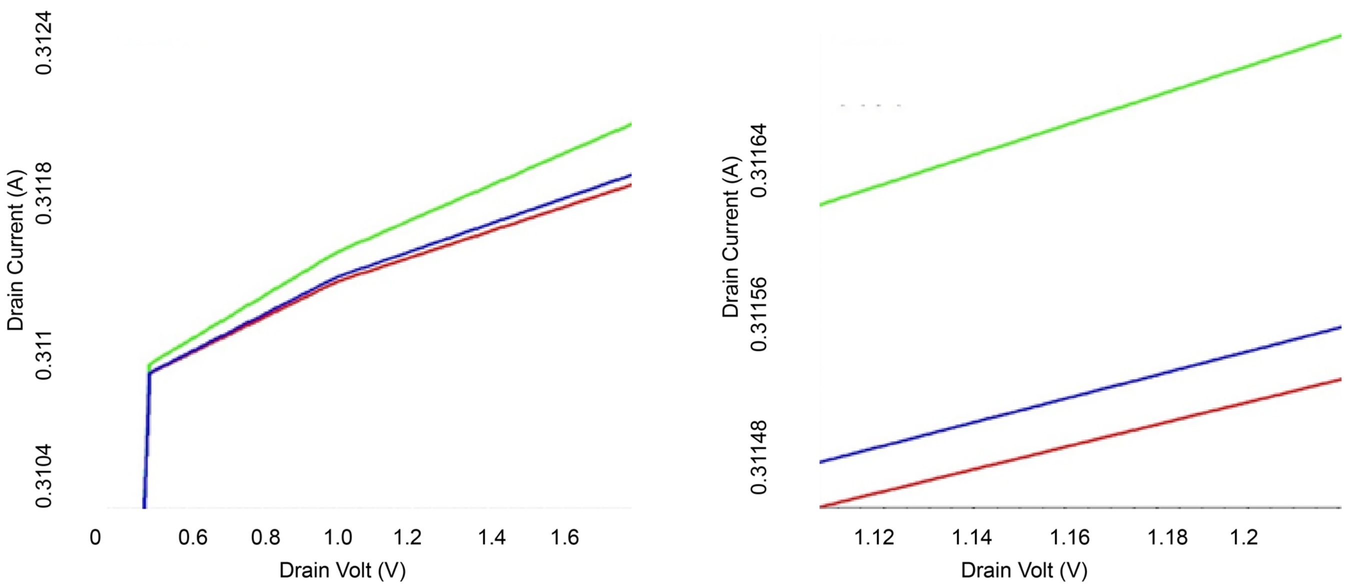



Figure 5. Change in drain current in the biosensor; the bottommost curve represents drain current of the clean device, the topmost curve represents the chemically modified device drain current, and middle curve represents the curve assuming an 87% conjugation success rate.

The device parameters associated with the developed clean device model can be seen in Table 1.

Table 2 shows the device physical characteristics at each step of its operation on different substrates.

The simulation results show that the polarizations remain almost constant at the interface regardless of the process step, and the substrate used. Also, as expected the charge concentration as well as the quantum well depth increases due to the chemically modified surface.

## 5. Experimental Results

Before any chemical modifications, the DC currents and voltages are measured using a DC probe station in conjunction with IC-CAP software. Figure 7 shows the ID—$V_{DS}$

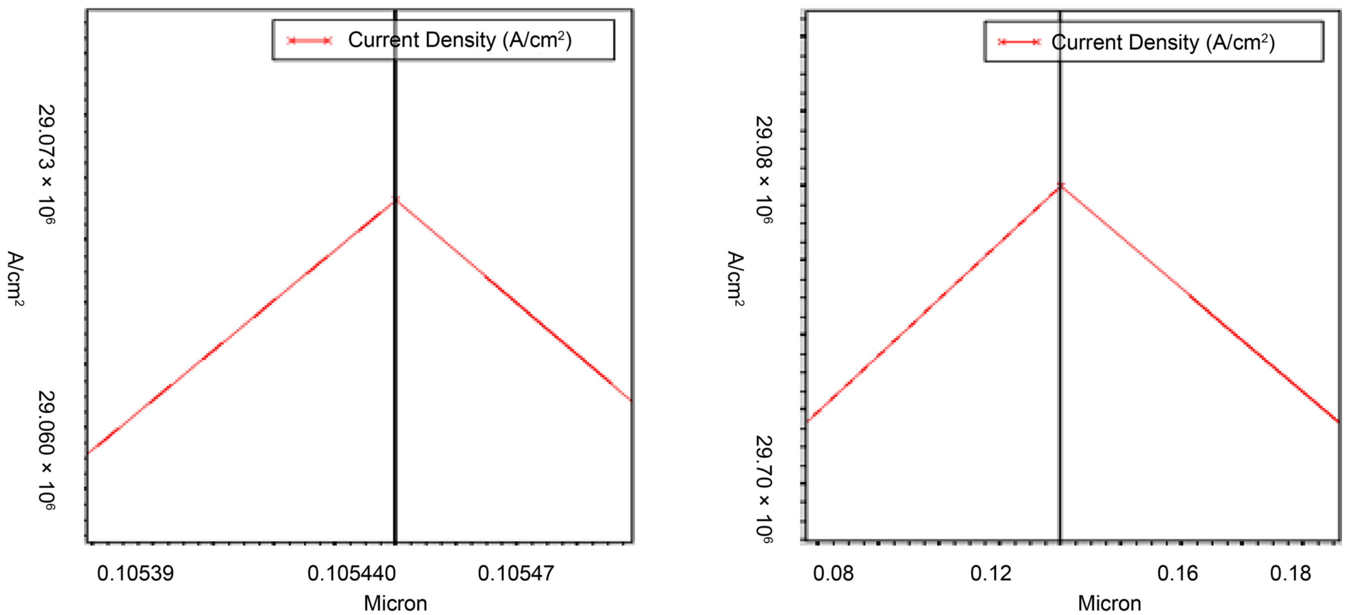



Figure 6. A change in current density at the interface is observed; Electron current density of device at the interface before conjugation (left) and Electron current density of device at the interface after conjugation (right).

Table 1. Modeled intrinsic device parameters taken at the interface.

| Parameter | Value |
|---|---|
| Charge Concentration (C/cm$^3$) | 0.75 |
| Cond. Current Density (A/cm$^2$) | 6.5e8 |
| Quantum Well Depth (eV) | 0.04 |
| Electric Field (V/cm) | 9e5 |
| Polarization Charge (C/cm$^3$) | 0.75 |
| Potential (V) | 1.9 |
| Mobility (x) (cm$^2$/V-S) | 920,000 |

**Table 2.** Physical characteristics of simulated biosensor taken at interface.

| Mode | Parameter | SiC | Sapphire | Diamond |
|---|---|---|---|---|
| Operating/Floating | Charge Concentration (C/cm$^3$) | 0.82246 | 0.8273 | 0.82311 |
| Operating/Floating | Cond. Current Density (A/cm$^2$) | 2.907e7 | 2.9067e7 | 2.9073e7 |
| Operating/Floating | Electric Field (V/cm) | 1.023e6 | 1.0245e6 | 1.0213e6 |
| Operating/Floating | Polarization Charge (C/cm$^3$) | 0.768 | 0.768 | 0.768 |
| Anti-MIG | Charge Concentration (C/cm$^3$) | 0.82522 | 0.8 | 0.82578 |
| Anti-MIG | Cond. Current Density (A/cm$^2$) | 2.907e7 | 2.9077e7 | 2.9075e7 |
| Anti-MIG | Electric Field (V/cm) | 1.02235e6 | 1.019e6 | 1.02055e6 |
| Anti-MIG | Polarization Charge (C/cm$^3$) | 0.768 | 0.768 | 0.768 |
| MIG | Charge Concentration (C/cm$^3$) | 0.82453 | 0.82613 | 0.82365 |
| MIG | Cond. Current Density (A/cm$^2$) | 2.9065e7 | 2.90685e7 | 2.90737e7 |
| MIG | Electric Field (V/cm) | 1.0247e6 | 1.0233e6 | 1.02043e6 |
| MIG | Polarization Charge (C/cm$^3$) | 0.768 | 0.768 | 0.768 |

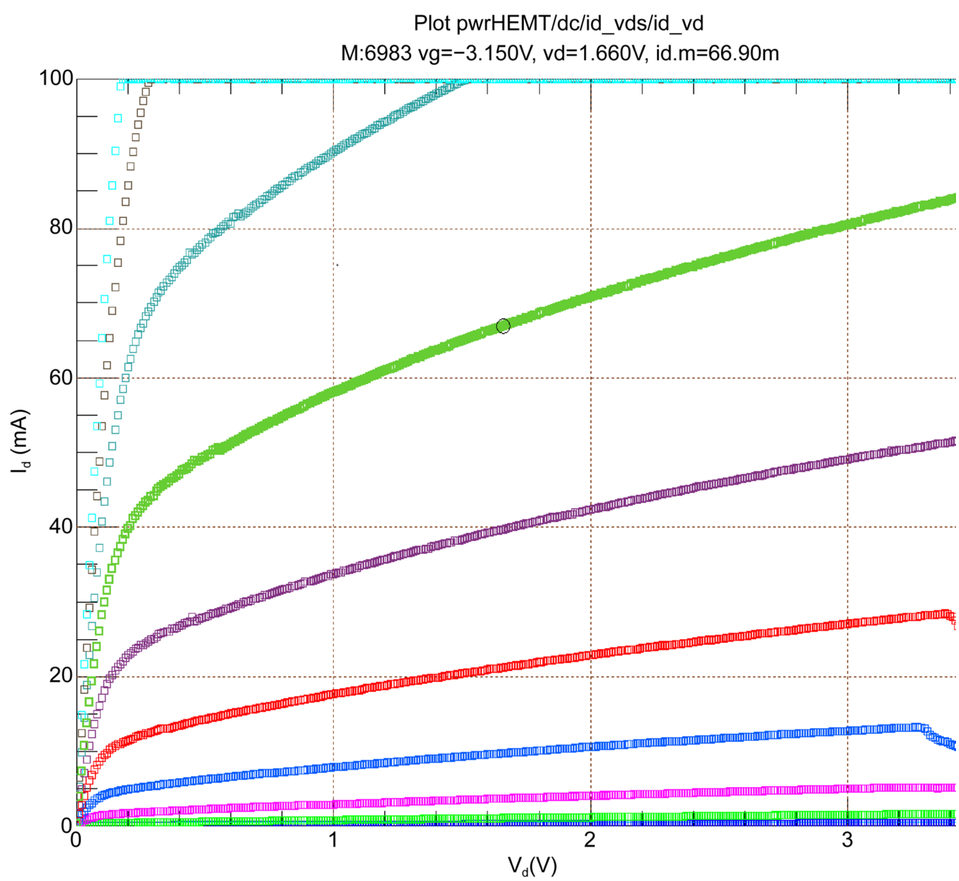



**Figure 7.** $I_d$ vs $V_{DS}$ curves at different gate voltages.

characteristics of the AlGaN/GaN HEMT. The knee voltages of about 0.2 - 0.4 V are observed throughout the multiple devices of the same type. Current collapse phenomenon is observed at higher $V_{DS}$ values. It is known that current collapse phenomenon occurs in AlGaN HEMT devices under AC and pulsed conditions [37]-[41].

To prepare the proper sensing environment, 1.826 g of 0.1 M phosphate buffer solution (PBS) is dissolved in 8mL of de-ionized water. A pH meter is used to verify a pH of 7.44. This is an environment which closely mimics the pH of blood. Then the cross-linker DSP is dissolved in 1 mL of organic solvent (Dimethyl Sulfoxide [DMSO]) to create an aqueous solution that could be used to coat the surface of the gate electrode on the HEMT device. After coating the device thoroughly, the device is incubated for 20 minutes at room temperature. The surface is rinsed with PBS to remove the surface of any unbinded DSP. After rinsing, the Anti-MIG is applied to the surface. The antibody is added to 0.1 mL of PBS with the appropriate pH. This step needs to be performed immediately after the incubation period to ensure proper protein coupling [25]. The device is then left to incubate at room temperature for 2 hours. Conjugation between proteins does not advance significantly after the first 1 - 2 hours, but incubation can be variable anywhere between 1 - 4 hours [25]. The gate surface is rinsed to remove unconjugated proteins, and the drain current under DC conditions are observed and compared to clean device operations. Upon conclusion that conjugation is successful, a small sample of MIG is added to PBS (with the appropriate pH) and introduced to the gate surface of the device, and the immediate response is observed in real time.

A simple circuit is constructed to observe real-time response of the device. In a floating gate configuration steady state current of the clean device is observed to be 66.89 mA. Upon construction of the SAM layer, an 80 μA rise in current is observed, which ensures that successful induction of electrons are exhibited. Upon introduction of a 2 μL sample of MIG in a 0.1M PBS onto the gate surface, a rapid response is observed with a 70 μA decrease in steady state current after 30 seconds. Clearer results are seen after 1 minute as demonstrated in **Figure 8**. The simulated results and experimental results are in good agreement with the approach taken. Small differences between the simulated and experimental results are observed due to trapping effects not taken into account by the 2-D simulation. Furthermore, the assumption that all charges in the MIG and Anti-MIG are live and equally distributed further contributes to the deviation.

## 6. Conclusions

A biosensor for the detection of Human MIG (CXCL9) by amperometric method is successfully demonstrated using AlGaN/GaN HEMT devices. A systematic approach is taken to create an improved two-dimensional model to investigate the quantum behavior upon chemical modification of the gate electrodes. The results seen in this research indicate that it is possible to build a reliable, chemically inert, and thermally stable biosensor.

The developed analytical model can be improved by addressing trapping and all parasitic effects in the two-dimensional numerical simulation. Further work needs to be

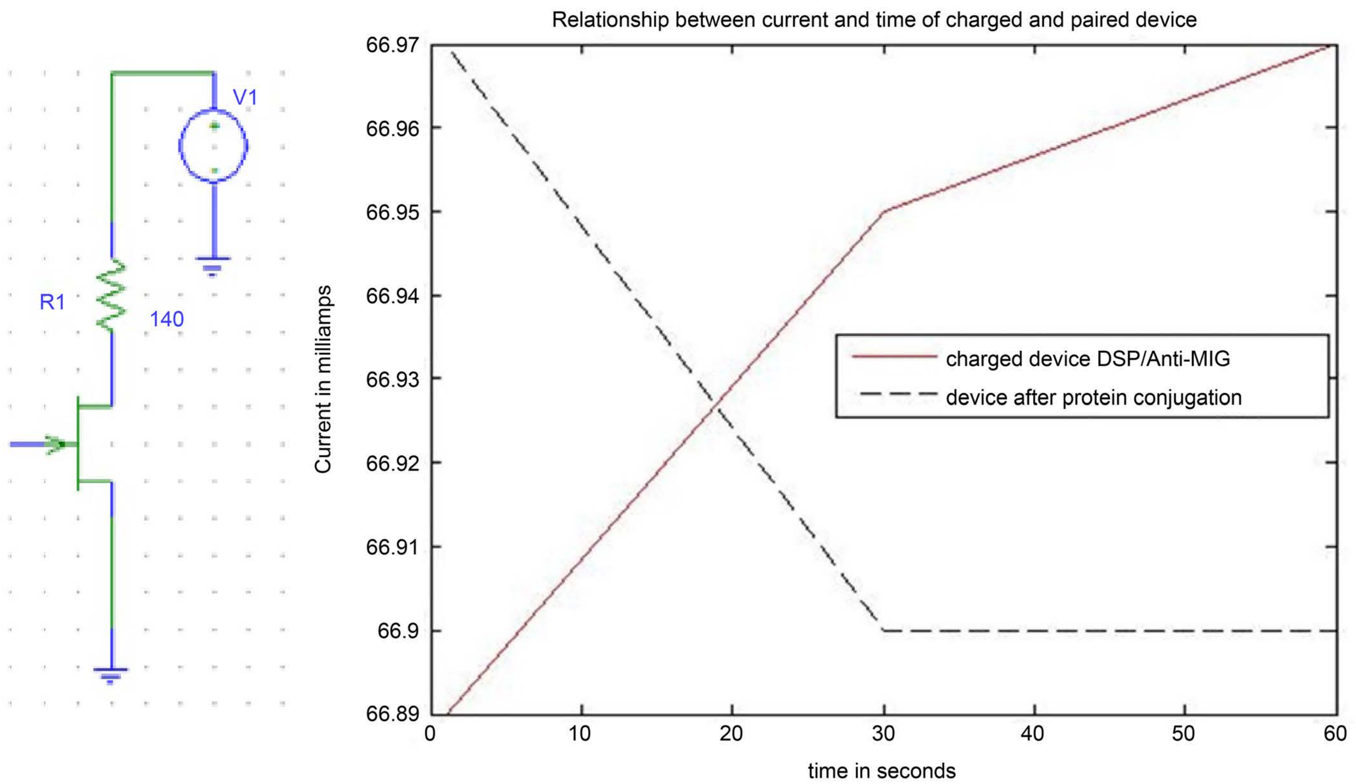



**Figure 8.** Change in current and time before and after introduction of MIG to previously prepared SAM layer. Results are comparable to steady state current of clean device (66.89 mA). Note: $V_1$ is 10 volts.

conducted to determine the repeatability of the biosensor. A positive shifting in threshold voltage has been observed after re-testing of a previously chemically modified device. Furthermore, the role of different type of substrates also needs to be investigated. 2-D simulation shows an improvement in performance of the HEMT device on a Sapphire substrate; however experimental work is needed to validate these results.

## Acknowledgements

This project is supported in part by the National Science Foundation (NSF) under grant Number: 0820381.

## References

[1] Tulip, F.S., Eteshola, E., Islam, S.K., Mostafa, S. and Huq, H.F. (2011) Label Free Detection of Human MIG Using AlGaN/GaN High Electron Mobility Transistor. *ISDRS* 2011, College Park, 7-9 December 2011. http://dx.doi.org/10.1109/isdrs.2011.6135242

[2] Kang, B.S., Ren, F., Wang, L., Lofton, C., Tan, W.W., Pearton, S.J., Dabiran, A., Osinsky, A., and Chow, P.P. (2005) Electrical Detection of Immobilized Proteins with Ungated AlGaN/GaN High-Electron-Mobility Transistors. American Institute of Physics, College Park.

[3] Eliza, S.A., Islam, S.K., Mostafa, S. and Tulip, F.S. (2010) Modeling of AlGaN/GaN HEMT Based Stress Sensors. 6*th International Conference on Electrical and Computer Engineering* (*ICECE*). http://dx.doi.org/10.1109/ICELCE.2010.5700689

[4] Tulip, F.S., Mostafa, S., Islam, S.K., Eteshola, E., Eliza, S.A., Lee, I., Greenbaum, E. and Evans, B.R. (2010) GaN-AlGaN High Electron Mobility Transistors for Multiple Biomolecule Detection Such as Photosystem I and Human MIG. 6*th International Conference on Electrical and Computer Engineering* (*ICECE*). http://dx.doi.org/10.1109/ICELCE.2010.5700690

[5] Wang, Y.-L., Huang, C.-C., Hsu, Y.-R. and Kang, Y.-W. (2013) Identification of Ligand-Receptor Binding Affinity Using AlGaN/GaN High Electron Mobility Transistors and Binding-Site Models. *NEMS* 2013, Suzhou, 7-10 April 2013, 532-538.

[6] Podolska, A., Seeber, R.M., Mishra, U.K., Pfleger, K.D.G., Parish, G. and Nener, B.D. (2012) Detection of Biological Reactions by AlGaN/GaN Biosensor. *Sensors and Actuators B Chemical*, **177**, 577-582.

[7] Salaman, J.R (2013) Monitoring of Rejection in Renal Transplantation. NCBI US National Library of Medicine, Bethesda.

[8] Andersen, C.B (2013) Acute Kidney Graft Rejection Morphology and Immunology. NCBI. US National Library of Medicine, Bethesda

[9] Bean, J.C. (1990) Materials and Technologies. In: Sze, S.M., Ed., *High-Speed Semiconductor Devices,* John Wiley & Sons, New York.

[10] Lenka, T.R. and Panda, A.K. (2011) Characteristics Study of 2DEG Transport Properties of AlGaN/GaN and AlGaAs/GaAs-Based HEMT. National Institute of Science and Technology, Brahmapur. http://journals.ioffe.ru/ftp/2011/05/p660-665.pdf

[11] Alur, S., Gnanaprakasa, T., Xu, H., Wang, Y.Q., Simonian, A.L., Oyarzabal, O.A. and Park, M. (2009) A Biosensor Based on GaN Field Effect Transistor. *CS MANTECH Conference*, Tampa, 18-21 May 2009.

[12] BioLegend (2013) Recombinant Human CXCL9 (MIG) (Carrier-Free) Technical Data Sheet. BioLegend. http://www.biolegend.com/recombinant-human-cxcl9-mig-carrier-free-7464.html

[13] Ebioscience (2013) Anti-Mouse CXCL9 (MIG) PE Technical Datasheet. Mouse CXCL9 (MIG) Antibody PE MIG-2F5.5 RUO. Affymetrix Ebioscience. http://www.ebioscience.com/mouse-cxcl9-antibody-pe-mig-2f55.htm

[14] AlphaLISA Research Reagents (2013) Human C-X-C Motif Chemokin 9/Monokine Induced by Gamma Interferon (CXCL9/MIG) Kit Tehnical Data Sheet. PerkinElmer. http://www.perkinelmer.com/Content/TDLotSheet/TDS-AL280-06%20CXCL9_MIG.pdf

[15] Trotta, T., Costantini, S. and Colonna, G. (2013) Modelling of the Membrane Receptor CXCR3 and Its Complexes with CXCL9, CXCL10, and CXCL11 Chemokines: Putative Target for New Drug Design. National Center for Biotechnology Information. US National Library of Medicine, Bethesda. http://www.ncbi.nlm.nih.gov/pubmed/19800124

[16] Groom, J.R. and Luster, A.D. (2011) *Experimental Cell Research*, **317**, 620-631. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3065205/ http://dx.doi.org/10.1016/j.yexcr.2010.12.017

[17] Mandal, A. (2013) Chemokine Receptors. News Medical. http://www.news-medical.net/health/Chemokine-Receptors.aspx

[18] SIno Biological Inc (2013) What Are Chemokines. http://www.sinobiological.com/What-is-Chemokine-a-5879.html

[19] Hu, H.Z., Kanmaz, T., Feng, P., Torrealba, J., Kwun, J., Fechner, J.H., Schultz, J.M., Dong, Y.C., Kim, H.T., Dar, W., Hamawy, M.M. and Knechtle, S.J. (2004) *Transplantation*, **78**, 613-614. http://dx.doi.org/10.1097/00007890-200407271-01651

[20] Whiting, D., Hsieh, G., Yun, J.J., Banerji, A., Yao, W., Fishbein, M.C., Belperio, J., Strieter, R.M., Bonavida, B. and Ardehali, A. (2004) *The Journal of Immunology*, **172**, 7417-7424. http://dx.doi.org/10.4049/jimmunol.172.12.7417

[21] Yun, J.J., Fischbein, M.P., Whiting, D., Irie, Y., Fishbein, M.C., Burdick, M.D., Belperio, J., Strieter, R.M., Laks, H., Berliner, J.A. and Ardehali, A. (2002) *The American Journal of Pathology*, **161**, 1307-1313. http://dx.doi.org/10.1016/S0002-9440(10)64407-0

[22] Tulip, F.S., Eteshola, E., Desai, S., Mostafa, S., Roopa, S. and Syed Islam, K. (2013) *IEEE Sensors Journal*, **13**, 438-439. http://dx.doi.org/10.1109/JSEN.2012.2226573

[23] Lee, S.C. (2016) Protein Engineering to Solve Problems at Nanoscale. Department of Biomedical Engineering, Department of Chemical and Biomolecular Engineering Ohio State University.

[24] AVIVA Systems Biology (2012) CXCL9 Antibody Used to Evaluate Immunofet Feasibility in Physiological Salt Environments.

[25] Thermo Scientific (2008) Tech Tip #2: Attach a Protein onto a Gold Surface. http://www.piercenet.com/files/TR0002-Attach-to-gold.pdf

[26] Shun, H.H., Lin, C.W. and Chiu, H.C. (2008) High Sensitivity PH Sensor Using AlGaN/GaN HEMT Heterostructure Design. *IEEE International Conference on Electron Devices and Solid-State Circuits*, Hong Kong, 8-10 December 2008.

[27] Eliza, S.A., Islam, S.K., Lee, I., Greenbaum, E. and Tulip, F.S. (2009) Modeling of a Floating Gate AlGaN/GaN Heterostructure-Transistor Based Sensor. *Semiconductor Device Research Symposium* (*ISDRS*'09), College Park, 9-11 December 2009.

[28] Charfeddine, M., Belmabrouk, H., Ali Zaidi, M. and Maaref, H. (2012) *Journal of Modern Physics*, **3**, 881-886. http://dx.doi.org/10.4236/jmp.2012.38115

[29] Rashmi, A. Haldar, K., Gupta, S.M. and Gupta, R.S. (2002) *Solid-State Electronics*, **46**, 621-630. http://dx.doi.org/10.1016/S0038-1101(01)00332-X

[30] Gangwani, P., Pandey, S., Haldar, S., Gupta, M. and Gupta, R.S. (2007) *Solid-State Electronics*, **51**, 130-135. http://dx.doi.org/10.1016/j.sse.2006.11.002

[31] Khir, F.L.M., Matthew, M., Anna, P., *et al.* (2014) *Applied Surface Science*, **314**, 850-857. http://dx.doi.org/10.1016/j.apsusc.2014.07.002

[32] Albrecht, J.D., Wang, R.P., Ruden, P.P., Farahmand, M. and Brennan, K.F. (1998) *Journal of Applied Physics*, **83**, 4777-4781. http://dx.doi.org/10.1063/1.367269

[33] Proteins, Enzymes and Peptides (2013) CXCL9 Recombinant Human Protein.

[34] Recombinant Human CXCL9 MIG Carrier-Free—BioLegend (2013)

[35] Recombinant Mouse CXCL9/MIG (2013)

[36] Abou El-Ela, F.M. and Mohamed, A.Z. (2013) *ISRN Condensed Matter Physics*, **2013**, Article ID: 654752.

[37] Albrecht, J.D., Ruden, P.P., Bellotti, E. and Brennan, K.F. *MRS Internet Journal of Nitride Semiconductor Research*, **4**.

[38] Xu, D., Chu, K., Diaz, J., Zhu, W., Roy, R., Mt Pleasant, L., Nichols, K., Chao, P.C., Xu, M. and Ye, P.D. (2013) *IEEE Electron Device Letters*, **34**, 744-746. http://dx.doi.org/10.1109/LED.2013.2255257

[39] Mishra, U.K., Shen, L., Kazior, T.E. and Wu, Y.F. (2008) *Proceedings of the IEEE*, **96**, 287-305. http://dx.doi.org/10.1109/JPROC.2007.911060

[40] Feng, Z., Xie, S., Zhou, R., Yin, J., Zhou, W. and Cai, S. (2010) *Journal of Semiconductors*, **31**, Article ID: 084001.

[41] Tanuma, N., Tacano, M., Yagi, S. and Sikula, J. (2009) *Journal of Statistical Mechanics: Theory and Experiment*, Article ID: P01038.
http://dx.doi.org/10.1088/1742-5468/2009/01/P01038

Journal of Modern Physics, 2016, 7, 1725-1737
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Time of the Energy Emission in the Hydrogen Atom and Its Electrodynamical Background

**Stanisław Olszewski**

Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, Poland
Email: olsz@ichf.edu.pl

**How to cite this paper:** Olszewski, S. (2016) Time of the Energy Emission in the Hydrogen Atom and Its Electrodynamical Background. *Journal of Modern Physics*, **7**, 1725-1737.
http://dx.doi.org/10.4236/jmp.2016.713155

**Received:** August 18, 2016
**Accepted:** September 25, 2016
**Published:** September 28, 2016



## Abstract

The time of the energy emission between two neighbouring electron levels in the hydrogen atom has been calculated first on the basis of the quantum aspects of the Joule-Lenz law, next this time is approached with the aid of the electrodynamical parameters characteristic for the electron motion in the atom. Both methods indicate a similar result, namely that the time of emission is close to the time period of the electromagnetic wave produced in course of the emission. As a by-product of calculations, the formula representing the radius of the electron microparticle is obtained from a simple combination of the expressions for the Bohr magnetic moment and a quantum of the magnetic flux.

## Keywords

Energy Emission in the Hydrogen Atom, Time of the Electron Transition between Two Quantum Levels, Electrodynamical Parameters Characteristic for the Electron Transition

## 1. Introduction

In physics, we often look for a simple explanation of the important phenomena without going much into details of the examined process. A well known example is the energy spectrum of the hydrogen atom. The first step to approach this spectrum theoretically was based on the idea that the force of the electrostatic attraction existent between the electron particle and the atomic nucleus remained in an equilibrium with the centrifugal force due to the circular electron motion about the same nucleus [1]. The second decisive step was that the angular momentum which accompanied the motion leading to the equilibrium of the atomic system should be quantized in a proper way. A combination of these two steps gave a spectacular success of the Bohr atomic model expressed in terms of positions of the electron energy levels present in the atom.

But the point not examined in the Bohr theory was the transition time between different quantum levels necessary to obtain the energy spectrum effect. A kind of paradox becomes that if we have two quantum states in the atom, say a hydrogen atom, their energies are well known, and the same knowledge applies naturally to the energy difference

$$\Delta E = E_{n+1} - E_n > 0 \tag{1}$$

between the states $n + 1$ and $n$ entering the transition process, but we cannot answer how long is the time interval $\Delta t$ necessary to perform such transition. Certainly, the interval $\Delta t$ is classified as "short" but no quasi-definite answer on its size is in practice available.

A reason of such situation cannot be the so-called uncertainty principle between the intervals of energy and time introduced by Heisenberg [2] [3]. In fact, this principle concerns rather a mutual relation between two definite and accessible intervals $\Delta E$ and $\Delta t$ entering a given quantum process than an "uncertainty" of the accuracy which can be attained in the measurement of the sizes of the mentioned intervals [4]-[6]. This implies that there does not exist an a priori difficulty to obtain $\Delta t$ when $\Delta E$ is known.

The main source of difficulty to calculate $\Delta t$ seems to be a probabilistic- and statistical character of examination applied in the treatment of the electron transitions. This kind of approach, being typical for the old quantum theory [7] [8], obtained its farther background in the formalism of quantum mechanics [3] [9] [10]. In effect the results for $\Delta t$ connected with the electron transition obtained respectively by the classical and quantum-mechanical approaches became diametrically different [10]. For, in order to obtain an agreement with the transition intensity of energy provided by the quantum-mechanical theory, the classical approach to that intensity required the time interval $\Delta t$ of an infinite size, viz.

$$\Delta t \to \infty, \tag{2}$$

instead of a finite (small) $\Delta t$ dictated evidently by an experimental practice.

In a set of papers [11]-[15], we tried to approach the size of $\Delta t$ with the aid of an examination of the electron transitions in small quantum systems with the aid of the Joule-Lenz law; see e.g. [16] [17]. For transitions connected with the population change of the neighbouring energy levels, *i.e.* $n + 1$ and $n$, the main result was that the relation

$$\Delta E \Delta t = h \tag{3}$$

should be satisfied. In the case of the harmonic oscillator

$$\Delta E = \hbar\omega = h\nu \tag{4}$$

where

$$\nu = \frac{1}{T_{\text{osc}}} \tag{5}$$

is the oscillator frequency of the emitted electromagnetic wave and $T_{\text{osc}}$ is the time period of that wave. In effect we obtain with the aid of (3)-(5):

$$h\nu\Delta t = \frac{h}{T_{\text{osc}}}\Delta t = h \tag{6}$$

or

$$\Delta t = T_{\text{osc}}. \tag{7}$$

Because of (3) the emission intensity of a single wave having the frequency $\nu$ becomes

$$\frac{\Delta E}{\Delta t} = \frac{(\Delta E)^2}{h} = \frac{(\hbar\omega)^2}{h} = \frac{h\omega^2}{(2\pi)^2}. \tag{8}$$

The intensity (8) can be referred to the the quantum-mechanical expression for the transition intensity of the harmonic oscillator [10]

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{2}{3}\frac{e^2\omega^3 n\hbar}{c^3 m} \tag{9}$$

by the formula

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\Delta E}{\Delta t}\gamma\frac{2\pi}{\omega}n = \frac{\Delta E}{\Delta t}\gamma T_{\text{osc}} n \tag{10}$$

where

$$\gamma = \frac{2}{3}\frac{e^2\omega^2}{mc^3} \tag{11}$$

is the damping term of the oscillator and $T_{\text{osc}}$ is the time period presented in (5); see also [10].

## 2. The Aim of the Paper

The aim of the present paper is to examine in some detail the transition time $\Delta t$ between the neighbouring quantum levels of the hydrogen atom. Certainly the size of $\Delta t$, because of its expected very short duration, seems to be hardly possible to be compared accurately with the experimental data. Nevertheless, an idea how $\Delta t$ can be influenced by the electrodynamical parameters responsible for the electron transition could be given. This makes, in principle, the problem of the time transfer $\Delta t$ between two quantum states reduced to a semiclassical one, so it can be treated with the aid of the classical electrodynamics. Before the electrodynamical properties will be discussed it seems of use to get an insight into the quantum aspects of $\Delta t$ based on the Joule-Lenz law.

## 3. Electron Transition Time Obtained from the Joule-Lenz Law

A preliminary approach to the time transfer of energy in the hydrogen atom, but not only in such system, can be done with the aid of a quantum insight into the Joule-Lenz law; see [11]-[15]. The dissipation rate of the energy $\Delta E$ within the time interval $\Delta t$ can be expressed by the formula

$$\frac{\Delta E}{\Delta t} = Ri^2. \tag{12}$$

Here the $R$ and $i$ are respectively the resistance and current intensity of the electron transition process done in course of $\Delta t$. The process leads to the energy dispense (decrease) equal to $\Delta E$.

For an unmodified (stationary) electron motion in the atom the current intensity is a constant

$$i_n = \frac{e}{T_n} \tag{13}$$

for any quantum state $n$ because of a constant time period $T_n$ representing the circulation of the electron charge $e$ about the nucleus [18]:

$$T_n = \frac{2\pi n^3 \hbar^3}{me^4}. \tag{14}$$

Certainly in course of the transition from one orbit, say that of the quantum state $n$ + 1, to the orbit of state $n$, the intensity is modified from $i_{n+1}$ to $i_n$, but in fact the change of $i$ is not large, especially when $n \gg 1$. Therefore, we can assume that

$$i = i_{n+1} = \frac{e}{T_{n+1}} \approx i_n = \frac{e}{T_n}. \tag{15}$$

The potential $V$ entering the electric resistance

$$R = \frac{V}{i} \tag{16}$$

let be

$$V = V_n = \frac{\Delta E}{e} \tag{17}$$

where in the numerator is substituted the energy difference between quantum states presented in (1). For large $n$, the energy $\Delta E$ becomes [18]:

$$\Delta E = E_{n+1} - E_n = \frac{me^4}{2\hbar^2}\left[\frac{1}{n^2} - \frac{1}{(n+1)^2}\right] \cong \frac{me^4}{\hbar^2 n^3}. \tag{18}$$

By combining $V_n$ in (17) with $i_n$ calculated from (13) and (14), we obtain the resistance

$$R = \frac{\Delta E}{ei_n} = \frac{me^4}{\hbar^2 n^3}\frac{1}{e^2}\frac{2\pi n^3 \hbar^3}{me^4} = \frac{2\pi\hbar}{e^2} = \frac{h}{e^2} \tag{19}$$

independent of the state $n$. The resistance $R$ obtained in (19) is characteristic for the integer quantum Hall effect [19]. The whole fraction (12) becomes

$$\frac{\Delta E}{\Delta t} = \frac{h}{e^2} i_n^2 = \frac{h}{e^2}\frac{e^2}{\left(2\pi n^3 \hbar^3\right)^2} m^2 e^8 = \frac{h}{e^2}\frac{e^{10} m^2}{(2\pi)^2 n^6 \hbar^6} = \frac{me^4}{\hbar^2 n^3}\frac{1}{\Delta t} \tag{20}$$

where in the last step we applied for $\Delta E$ the result of (18) obtained from large $n$.

In effect, $\Delta t$ attains the value

$$\Delta t = \frac{me^4}{\hbar^2 n^3}\frac{2\pi}{e^8 \hbar}\frac{\hbar^6 n^6}{m^2} = \frac{2\pi \hbar^3 n^3}{e^4 m} \tag{21}$$

which is formally equal to the size of the time period $T_n$ given in (14).

In the next step, by multiplying the both sides of (21) by $\Delta E$ taken from the final step of (18) we obtain

$$\Delta E \Delta t = \frac{me^4}{\hbar^2 n^3} \frac{2\pi\hbar^3 n^3}{me^4} = 2\pi\hbar = h. \tag{22}$$

This is a result providing us with a very simple relation between the $\Delta E$ and $\Delta t$.

The relation identical with (3) has been rather extensively applied in comparing the quantum-mechanical spectrum of transition probabilities between the electron states in the hydrogen atom [20] [21] with the intensity of the electron transitions calculated with the aid of $\Delta E$ and $\Delta t$ entering the formula (22), *i.e.* the formula

$$\frac{\Delta E}{\Delta t} = \frac{(\Delta E)^2}{h} \tag{23}$$

similar to (8) for the oscillator has been applied; see [11]-[15]. More precisely, the formula (23) is valid solely for transitions

$$n+1 \to n, \tag{24}$$

but the transition time $\Delta t$ corresponding, say, to situations

$$\begin{aligned} &n+2 \to n, \\ &n+3 \to n, \\ &n+4 \to n, \end{aligned} \tag{25}$$

etc, can be composed from the $\Delta t$ calculated for the case of (23) [13] [14].

Another important point concerning $\Delta t$ in (23) is its reference to the time period $T$ of the electromagnetic wave produced by the energy difference $\Delta E$. In fact because of the result

$$\Delta t = T_n \tag{21a}$$

obtained in (21), the formula (23) becomes reduced to

$$\frac{1}{\Delta t} = \frac{\Delta E}{h} = \nu = \frac{1}{T}, \tag{26}$$

therefore we obtain

$$\Delta t = T = T_n \tag{27}$$

which is similar to (7) for the harmonic oscillator.

In fact, the formulae (23) and (27) are not specific solely for the hydrogen atom and the harmonic oscillator, but their validity can be extended to other quantum systems, for example the particle in a one-dimensional potential box; see [11] [12] [15]. Because of (21) and (27), we obtain also

$$\frac{h}{T_n} = \frac{h}{T} = h\nu = \Delta E. \tag{27a}$$

In fact on the basis of of (14), we have

$$\frac{h}{T_n} = \frac{hme^4}{2\pi n^3 \hbar^3} = \frac{me^4}{n^3 \hbar^2} = \Delta E. \tag{27b}$$

In the last step of (27b) the result obtained in (18) has been taken into account.

The aim of the remainder of the paper is to show—for the hydrogen atom taken as an example—that the result of (26) obtained mainly on a quantum footing—could find its correspondence also in effect of a semiclassical approach to the electron transition.

## 4. Electrodynamical Parameters Connected with the Electron Transition and Its Current

First of the necessary parameters will be the magnetic flux $\Phi$ and its changes in the atom due to the changes of the population of the quantum levels [15]. The magnetic field $B_n$ connected with the quantum state $n$ is an effect of the electron circulation along the orbit $n$. This implies that $T_n$ in (14) and the time period causing the existence of $B_n$ should be equal. Therefore (see e.g. [22] [23])

$$\Omega_n = \frac{2\pi}{T_n} = \frac{eB_n}{mc} \tag{28}$$

where the field $\boldsymbol{B}$ is directed normally to the orbit plane.

A substitution of $T_n$ from (14) into (28) gives the relation

$$2\pi \left( \frac{2\pi n^3 \hbar^3}{me^4} \right)^{-1} = \frac{me^4}{n^3 \hbar^3} = \frac{eB_n}{mc} \tag{29}$$

from which

$$B_n = \frac{m^2 e^3 c}{n^3 \hbar^3}. \tag{30}$$

The magnetic flux across the area of a circular orbit having the radius [18]

$$r_n = \frac{n^2 \hbar^2}{me^2} \tag{31}$$

is equal to

$$\Phi_n = B_n S_n = \frac{m^2 e^3 c}{n^3 \hbar^3} \pi \frac{n^4 \hbar^4}{m^2 e^4} = \pi \frac{\hbar cn}{e} = \frac{hcn}{2e} \tag{32}$$

because the area enclosed by the orbit amounts to

$$S_n = \pi r_n^2 = \pi \left( \frac{n^2 \hbar^2}{me^2} \right)^2 = \pi \frac{n^4 \hbar^4}{m^2 e^4}, \tag{33}$$

on condition the radius $r_n$ in (31) is taken into account.

Evidently the absolute change of $\Phi_n$ associated with the change of $n$ by

$$|\Delta n| = 1 \tag{34}$$

provides us for any $n$ with the value

$$\Delta\Phi = \frac{hc}{2e}. \tag{35}$$

This is a quantum a) independent of $n$, b) well-known from the physics of superconductors [23]-[25]. Let us note that if

$$\mu_B = \frac{e\hbar}{2mc} \tag{35a}$$

is the Bohr magneton [3], the formula (35a) multiplied by $2\pi$ and divided by the flux in (35) gives

$$\frac{2\pi\mu_B}{\Delta\Phi} = 2\pi\frac{e\hbar}{2mc}\frac{2e}{hc} = \frac{e^2}{mc^2} \tag{35b}$$

which is a distance known as the radius of the electron microparticle; see e.g. [17].

In fact the steady orbital current $i_{n+1}$ is perturbed in course of transition from $n+1$ to $n$, nevertheless we expect this perturbation is small. The effective current of transition let be

$$i_{tr} = i_n + i_d(t) \tag{36}$$

where solely

$$i_d(t) \ll i_n \tag{37}$$

is the current part dependent on time $t$. Assuming that the orbits system between states $n+1$ and $n$ behaves like a condenser, our idea is to introduce a current

$$\frac{\mathrm{d}e}{\mathrm{d}t} = i_d \tag{38}$$

representing a discharge of the condenser [17] [26]. This

$$i_d(t) = i_d \tag{39}$$

enters (36) and (37).

The interval $\Delta\Phi$ in (35) is coupled with the self-induction constant $L$ by the formula

$$\frac{1}{c}\Delta\Phi = Li_{tr} = L(i + i_d), \tag{40}$$

but the differentiation process with respect to time concerns solely the term $i_d$:

$$\frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\Delta\Phi = L\frac{\mathrm{d}i_d}{\mathrm{d}t}. \tag{41}$$

The resistance $R$ is

$$R = \frac{V}{i} = \frac{\Delta E}{ei} \tag{42}$$

and the capacitance $C$ for a planar condenser is

$$C = \frac{e}{V} = \frac{e^2}{\Delta E}. \tag{43}$$

But because of a cylindrical shape of the orbits forming the condenser the formula (36) should be replaced by [26]

$$C = \frac{2e^2}{\Delta E}. \tag{44}$$

The $L$, $R$ and $C$ parameters enter the time-dependent differential equation for the current $i_d$ in (36) (see [17]):

$$L\frac{\mathrm{d}^2 i_d}{\mathrm{d}t^2}+R\frac{\mathrm{d}i_d}{\mathrm{d}t}+\frac{1}{C}i_d=\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_{\mathrm{appl}} \tag{45}$$

where $\mathcal{E}_{\mathrm{appl}}$ is the applied electromotive force. If we assume that

$$\mathcal{E}_{\mathrm{appl}}=\mathcal{E}_{\mathrm{oappl}}\mathrm{e}^{-j\omega t}, \tag{46}$$

where

$$j=(-1)^{1/2}, \tag{47}$$

the solution of (45) becomes

$$i_d=i_{d_0}\mathrm{e}^{-j\omega t}; \tag{48}$$

the $e$ in (46) and (48) is the basis of the natural logarithms.

On condition we assume that

$$\mathcal{E}_{\mathrm{appl}}=0, \tag{49}$$

the differential process of (45) performed upon (48) gives

$$\left(-L\omega^2-Rj\omega+\frac{1}{C}\right)i_d=0. \tag{50}$$

The solution of (50) leads to the frequency [17] [26]:

$$\omega=-j\frac{R}{2L}\pm\sqrt{\frac{1}{LC}-\left(\frac{R}{2L}\right)^2}. \tag{51}$$

## 5. Calculation of the Frequency $\omega$

From Equation (40), we have

$$L\cong\frac{\Delta\Phi}{ci}=\frac{h}{2ei} \tag{52}$$

because of (35), so in view of (44):

$$LC=\frac{\Delta\Phi}{ci}\frac{2e^2}{\Delta E}=\frac{h}{2ei}\frac{2e^2}{\Delta E}=\frac{eh}{\Delta Ei}. \tag{53}$$

In the next step from (42) and (52)

$$\frac{R}{2L}=\frac{1}{2}\frac{\Delta E}{ei}\frac{2ei}{h}=\frac{\Delta E}{h}. \tag{54}$$

Therefore, (51) becomes

$$\omega=-j\frac{\Delta E}{h}\pm\sqrt{\frac{i\Delta E}{eh}-\left(\frac{\Delta E}{h}\right)^2}=-j\frac{\Delta E}{h}\pm\sqrt{\frac{e}{T}\frac{\Delta E}{eh}-\left(\frac{\Delta E}{h}\right)^2}. \tag{55}$$

Since

$$h\nu=\frac{h}{T}=\Delta E \tag{56}$$

we have

$$\frac{1}{T} = \frac{\Delta E}{h}, \tag{57}$$

so the expression entering the square root in (55) is equal to

$$\frac{e\Delta E}{Teh} - \left(\frac{\Delta E}{h}\right)^2 = 0. \tag{58}$$

In effect $\omega$ in (51) becomes equal to the imaginary value

$$\omega = -j\frac{\Delta E}{h} = -j\frac{1}{T}. \tag{59}$$

This result gives for the current in (48) the formula

$$i_d = i_{d_0}\mathrm{e}^{-j\omega t} = i_{d_0}\mathrm{e}^{-\frac{t}{T}}, \tag{60}$$

With $e$ in (60) being the basis of natural logarithms, we obtain for $i_d$ a current exponentially decreasing with time $t$.

Let us note that when $R = 0$, which is the case where no resistance does exist for the transition current, the expression for $\omega$ becomes [see (51)]:

$$\omega\big|_{R=0} = \sqrt{\frac{1}{LC}} = \sqrt{\frac{\Delta E i}{eh}} = \sqrt{\frac{\Delta E}{Th}} = \frac{\Delta E}{h} = \frac{1}{T} = \nu \tag{61}$$

in virtue of (57).

## 6. Emission Rate and Its Damping Time

On the basis of (20) and (21), it is easy to calculate the emission rate

$$\frac{\Delta E}{\Delta t} = \frac{me^4}{\hbar^2 n^3}\frac{e^4 m}{2\pi\hbar^3 n^3} = \frac{m^2 e^8}{2\pi\hbar^5 n^6}. \tag{62}$$

By substituting for simplicity $m \approx \hbar \approx 10^{-27}$ and $e \approx 4.8\times10^{-10}$, we obtain

$$\frac{\Delta E}{\Delta t} \cong \frac{\left(10^{-27}\right)^{2-5}}{2\pi n^6}\left(4.8\times10^{-10}\right)^8 \approx \frac{10^6\times2.8}{2\pi n^6}\,\mathrm{erg/sec}. \tag{63}$$

This is a very high number especially for small $n$, nevertheless it is valid solely at the very beginning of the emission process. The duration of that process for the energy interval $\Delta E$ is approximately equal to [see (21)]:

$$\Delta t = \frac{\Delta E}{\Delta E/\Delta t} = \frac{2\pi\hbar^3 n^3}{e^4 m} \cong \frac{2\pi\times\left(10^{-27}\right)^{3-1} n^3}{\left(4.8\times10^{-10}\right)^4} \approx \frac{2\pi n^3}{5.3}10^{-16}\,\mathrm{sec}. \tag{64}$$

## 7. Velocity of the Electron Transition between Two Neighbouring Quantum Levels

The result of (21) and (26) allows us to calculate the velocity of transition of the electron particle between the levels $n + 1$ and $n$. This is

$$v_{tr} = \frac{\Delta r_n}{\Delta t} = \frac{\Delta r_n}{T_n} = \frac{\Delta E}{h}\Delta r_n, \tag{65}$$

where [see (31)]

$$\Delta r_n = \frac{(n+1)^2 - n^2}{me^2}\hbar^2. \tag{66}$$

With $\Delta E$ represented by a difference in the kinetic electron energy

$$\Delta E = -\Delta E_{\text{kin}} = \frac{m}{2}\left(v_n^2 - v_{n+1}^2\right) = \frac{m}{2}\left(\frac{e^2}{\hbar}\right)^2 \left[\frac{1}{n^2} - \frac{1}{(n+1)^2}\right],$$

this gives

$$v_{tr} = \frac{e^4}{h\hbar^2}\frac{m}{2}\frac{(n+1)^2 - n^2}{n^2(n+1)^2}\frac{(n+1)^2 - n^2}{me^2}\hbar^2 \cong \frac{e^2}{h}\frac{1}{2}\frac{(2n)^2}{n^4} = \frac{e^2}{\hbar}\frac{1}{\pi n^2} = v_n \frac{1}{\pi n}. \tag{67}$$

We find that the transition velocity of the electron between levels $n + 1$ and $n$ is by a factor of $1/\pi n$ smaller than the velocity $v_n$ along the orbit $n$. This calculation is fully original and new.

## 8. Comments

Heisenberg strongly criticized the Bohr atomic model as useless because it applied the unobserved elements of the atomic structure like the electron orbits; see e.g. [26]-[28].

Nevertheless the combined orbital parameters, like the orbit radius or orbit length and the time period of the electron circulation, allowed us to approach correctly the parts of the electron kinetic and potential energy which—when added together—gave a proper total electron energy in the atom. This energy formula has been next confirmed by both the modern quantum theory (quantum mechanics) and experiment.

But the modern theory did not provide us with an adequate information on the time duration of the electron processes in the atom, for example the time of the electron transitions. In this circumstance, a step towards the old quantum theory which applied definite periods of time seemed to be both realistic and useful.

In the first step, we assumed that the classical Joule-Lenz theory can couple the amount of energy emitted in the quantum process of an electron transition with the time necessary for that process. This assumption led to an extremely simple relation between the emitted energy $\Delta E$ and emission time $\Delta t$. In the present paper our aim was to approach the time necessary for the emission of $\Delta E$ on a somewhat different way than a direct application of the Joule-Lenz law, *i.e.* mainly with the aid of a classical analysis of the electric current produced as an effect of transition giving the energy change $\Delta E$.

In the first step of this analysis, the quanta of the magnetic induction and magnetic flux are introduced to the formalism. It should be noted that the Bohr magneton divided by the quantum change of the magnetic flux between the neighbouring levels [see (35b)] gives the well-known formula for the radius of the electron microparticle; see e.g. [17]. An earlier derivation of (35b)—different only in a constant factor—has been done in [29]; see also [30].

Next the electric current connected with the transition between two neighbouring quantum levels is considered as due to a discharge of the condenser. In order to examine this current, the costants of self-induction, resistance and capacitance characteristic for such condenser have been calculated. In effect an exponential decrease of the time-dependent part of the discharge current is obtained; see (60).

A very simple approach to the discharge current from state $n$ can be attained when the Ohm's law is applied [21]:

$$i_{on} = \frac{Q}{RC} = -\frac{\mathrm{d}Q}{\mathrm{d}t}. \tag{68}$$

Here $Q = e$ and $-\mathrm{d}Q = \mathrm{d}e$ is a small decrease of charge of the condenser in a small time interval $\mathrm{d}t$. The formula (68) gives

$$\frac{\mathrm{d}Q}{Q} = -(RC)^{-1}\,\mathrm{d}t. \tag{69}$$

Therefore

$$\log Q = -\frac{t}{RC} + \mathrm{cont} = -\frac{t}{2T} + \mathrm{cont} \tag{70}$$

because

$$RC = \frac{\Delta E}{ei}\frac{2e^2}{\Delta E} = \frac{2e}{i} = 2T. \tag{71}$$

Hence the charge $Q$ decreases with $t$ according to the formula

$$Q = \mathrm{e}^{-\frac{t}{2T}}. \tag{72}$$

The exponent of the natural logarithm basis e in (72) is a half of that obtained in (60). The rate of the emission in the form of the electromagnetic field energy has been discussed in [31].

It can be noted that Equations (3) and (22) are formally similar to the inequality proposed by Heisenberg called the uncertainty principle for energy and time. In fact the physical background for the intervals $\Delta E$ and $\Delta t$ entering he Heisenberg principle is much different than the properties of the intervals $\Delta p_x$ and $\Delta x$, concerning—for example—the $x$-coordinates of the momentum and position of a particle. For, contrary to the momenum and position, the energy $E$ can be measured to any degree of accuracy at any instant of time. Therefore $\Delta E$ can be the difference between two exactly measured values of energy at two different instants; see [4]-[6].

## 9. Conclusions

The paper approaches a seldom discussed problem of an individual electron transition between two quantum levels in the hydrogen atom. Consequently, no reference has been done to the well-known probabilistic theory usually applied to the quantum transitions.

In the first step, the emission time of energy between two neighbouring levels in the atom is calculated on the basis of the quantum aspects of the Joule-Lenz law; see (21) and (22). This time is found equal to the oscillation period of the electromagnetic wave emitted in course of the transition process; see (27).

Next, the problem of the emission was approached with the aid of the classical electrodynamics by assuming that the electron transition in the atom was roughly equivalent to a discharge of an electrical condenser. The damping time of the current obtained in course of such discharge is found to be close to the transition time attained in the Joule-Lenz theory; see (60) and (61), as well as (72) for the case of a simplified treatment of the calculation.

By assuming that the transition time between the quantum levels is similar to the emission time, the velocity of transition of the electron particle between the neighbouring orbits in the atom has been estimated; see (67).

A by-product of the calculations is the result that the Bohr magneton divided by the quantum of the magnetic flux obtained from the flux difference of two energy levels in the atom approaches the geometrical radius attributed to the electron microparticle; see (35a) and (35b).

## References

[1] Bohr, N. (1913) *Philosophical Magazine*, **26**, 1, 476, 827.

[2] Heisenberg, W. (1927) *Zeitschrift fuer Physik*, **43**, 172-198. http://dx.doi.org/10.1007/BF01397280

[3] Schiff, L.I. (1968) Quantum Mechanics. 3rd Edition, McGraw-Hill, New York.

[4] Landau, L. and Peierls, R. (1931) *Zeitschrift fuer Physik*, **69**, 56-69. http://dx.doi.org/10.1007/BF01391513

[5] Jammer, M. (1974) The Philosophy of Quantum Mechanics. Wiley, New York.

[6] Schommers, W. (1989) Space-Time and Quantum Phenomena, in “Quantum Theory and Pictures of Reality” ed. by W. Schommers. Springer, Berlin. http://dx.doi.org/10.1007/978-3-642-95570-9_5

[7] Planck, M. (1910) Acht Vorlesungen ueber theoretische Physik. S. Hirzel, Leipzig 1910.

[8] Einstein, A. (1917) *Physikalische Zeitschrift*, **18**, 121.

[9] Slater, J.C. (1960) Quantum Theory of the Atomic Structure. Vol. 1, McGraw-Hill, New York.

[10] Mizushima, M. (1970) Quantum Mechanics of Atomic Spectra and Atomic Structure. Benjamin, New York.

[11] Olszewski, S. (2015) *Journal of Modern Physics*, **6**, 1277-1288. http://dx.doi.org/10.4236/jmp.2015.69133

[12] Olszewski, S. (2016) *Journal of Modern Physics*, **7**, 162-174. http://dx.doi.org/10.4236/jmp.2016.71018

[13] Olszewski, S. (2016) *Journal of Modern Physics*, **7**, 827-851. http://dx.doi.org/10.4236/jmp.2016.78076

[14] Olszewski, S. (2016) *Journal of Modern Physics*, **7**, 1004-1020. http://dx.doi.org/10.4236/jmp.2016.79091

[15] Olszewski, S. (2016) *Reviews in Theoretical Science*, **4**, 336-352. http://dx.doi.org/10.1166/rits.2016.1066

[16] Lass, H. (1950) Vector and Tensor Analysis. McGraw-Hill, New York.

[17] Matveev, A.N. (1964) Electrodynamics and the Theory of Relativity. Izd. Wyzszaja Szkola, Moscow. (In Russian)

[18] Sommerfeld, A. (1931) Atombau und Spektrallinien. 5th Edition, Vol. 1, Vieweg, Braunschweig.

[19] MacDonald, A.H. (1989) Quantum Hall Effect: A Perspective. Kluwer, Milano.

[20] Bethe, H. (1933) Quantenmechanik der Ein- und Zwei-Elektronen Probleme. In: Geiger, H. and Scheel, K., Eds., *Handbuch der Physik*, Vol. 24, Part 1, Springer, Berlin.

[21] Condon, E.U. and Shortley, G.H. (1970) The Theory of Atomic Spectra. University Press, Cambridge.

[22] Slater, J.C. (1967) Quantum Theory of Molecules and Solids. Vol. 3, McGraw-Hill, New York.

[23] Kittel, C. (1987) Quantum Theory of Solids. 2nd Edition, Wiley, New York.

[24] Ziman, J.M. (1972) Principles of the Theory of Solids. 2nd Edition, University Press, Cambridge. http://dx.doi.org/10.1017/CBO9781139644075

[25] Cyrot, M. and Pavuna, D. (1992) Introduction to Superconductivity and High $T_c$ Materials. World Scientific, Singapore. http://dx.doi.org/10.1142/1039

[26] Jauncey, G.E.M. (1948) Modern Physics. 3rd Edition, Van Nostrand, New York.

[27] Van der Waerden, B.L. (Ed.) (1968) Discussion in "Sources of Quantum Mechanics". Dover Publications, New York.

[28] De Boer, J., Dal, E. and Ulfbeck, O. (Eds.) (1986) The Lesson of Quantum Theory. North-Holland, Amsterdam.

[29] Weyl, H. (1923) Raum-Zeit-Materie. Springer, Berlin. http://dx.doi.org/10.1007/978-3-642-98950-6

[30] Olszewski, S. (2016) *Journal of Modern Physics*, **7**, 1297-1303. http://dx.doi.org/10.4236/jmp.2016.711114

[31] Olszewski, S. *Journal of Modern Physics*, **7**, 1440-1448.

Journal of Modern Physics, 2016, 7, 1738-1748
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Affine Eikonal, Wavization and Wigner Function

**Akihiro Ogura**

Laboratory of Physics, Nihon University, Matsudo, Japan
Email: ogura.akihiro@nihon-u.ac.jp

**How to cite this paper:** Ogura, A. (2016) Affine Eikonal, Wavization and Wigner Function. *Journal of Modern Physics*, **7**, 1738-1748.
http://dx.doi.org/10.4236/jmp.2016.713156

**Received:** August 22, 2016
**Accepted:** September 25, 2016
**Published:** September 28, 2016



## Abstract

The aim in this paper is to construct an affine transformation using the classical physics analogy between the fields of optics and mechanics. Since optics and mechanics both have symplectic structures, the concept of optics can be replaced by that of mechanics and vice versa. We list the four types of eikonal (generating functions). We also introduce a unitary operator for the affine transformation. Using the unitary operator, the kernel (propagator) is calculated and the wavization (quantization) of the Gabor function is discussed. The dynamic properties of the affine transformed Wigner function are also discussed.

## Keywords

Affine Eikonal, Wavization of Gabor Function, Wigner Function

## 1. Introduction

Geometrical optics serves as a powerful tool for investigating optical systems. The path of a light ray is described by an eikonal. When the light rays are paraxial rays, this is classified as linear optics. In this approximation, the propagation of the light ray is described by the product of the refraction and the transfer matrices [1] [2]. Moreover, if the light ray is considered to have rotational symmetry with respect to the optical axis (skewness equal to zero), it is called meridian or Gaussian optics [3]. In this case, the refraction and transfer matrices are expressed by $2 \times 2$ matrices. The product of these matrices is also represented by a $2 \times 2$ matrix, and this is called an *ABCD*-matrix and specifies the optical system.

In general, the *ABCD*-matrix is specified by a three parameter ($A$, $B$, $C$, $D$ with $AD - BC = 1$) class of linear transformations [4] [5] in position and momentum. Linear canonical transformations have been studied by many authors at different times in different contexts. Good reviews can be found in [6] [7] and the references therein. Due to the condition $AD - BC = 1$, the *ABCD*-transformation is an area preserving transformation in phase space. Therefore, the Wigner function is only distorted in phase space but does not move in it.

In this article, we develop the mathematical properties of an affine transformation from the optical and mechanical points of view. Since the affine transformation has a displacement part, we are able to discuss the translation in phase space. Thus, we show that the affine transformation not only distorts but also displaces the Wigner function. Because this displacement can have time dependency, the Wigner function moves dynamically in phase space.

This paper is organized in the following way. In Section 2, we define the affine transformation and show the eikonals which generate this transformation. In Section 3, we turn to the quantum mechanical case for the affine transformation. We show that the operator of the affine transformation is obtained from the product of the displacement operator and the unitary operator of the *ABCD*-transformation. We also calculate the kernels of the affine transformation. In Section 4, we treat the wavization by referring to the Gabor function. In Section 5, we discuss the affine transformation of the Wigner function. We give an explicit form of the affine transformed Wigner function and examine the change in its configuration and the displacement of the Wigner function. Section 6 is devoted to a summary.

## 2. Affine Eikonal

The general affine transformation is defined by a linear combination of position $q$ and momentum $p$ with the four parameters $A$, $B$, $C$ and $D$ and the displacements for position $E$ and momentum $F$. We define the affine transformed position $Q$ and momentum $P$ as

$$Q = Aq + Bp + E, \tag{1a}$$

$$P = Cq + Dp + F, \tag{1b}$$

with the lossless (area-preserving or power-preserving) condition

$$AD - BC = 1. \tag{2}$$

In classical mechanics, this condition comes from which affine transformation $(Q, P)$ satisfies the Poisson bracket (PB)

$$\{Q, P\}_{PB} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial P}{\partial q}\frac{\partial Q}{\partial p} = 1, \tag{3}$$

that is, $Q$ and $P$ are canonical variables [8].

In geometrical optics, the path of the light ray is described by an eikonal. In the following discussion, we restrict ourselves to Gaussian optics, so each $q$ and $p$ is one-dimensional variable. There are four types of eikonal in Gaussian optics. We list the four types below;

$$V_1 = \frac{1}{2}(+QP - qp + FQ - EP) \quad \text{for point eiknal,} \tag{4a}$$

$$V_2 = \frac{1}{2}(-QP - qp + FQ - EP) \quad \text{for mixed eikonal,} \tag{4b}$$

$$V_3 = \frac{1}{2}(+QP + qp + FQ - EP) \quad \text{for mixed eikonal,} \tag{4c}$$

$$V_4 = \frac{1}{2}(-QP + qp + FQ - EP) \quad \text{for angle eikonal.} \tag{4d}$$

By substituting (1) into (4), we rewrite these eikonals in terms of two of the four canonical variables $q$, $p$, $Q$ and $P$,

$$V_1 = +\frac{A}{2B}q^2 + \frac{D}{2B}(Q-E)^2 - \left(\frac{q}{B} - F\right)(Q-E) + \frac{EF}{2} = -W_1(q,Q,t) \tag{5a}$$

$$V_2 = +\frac{C}{2D}q^2 - \frac{B}{2D}(P-F)^2 - \left(\frac{q}{D} + E\right)(P-F) - \frac{EF}{2} = -W_2(q,P,t) \tag{5b}$$

$$V_3 = -\frac{B}{2A}p^2 + \frac{C}{2A}(Q-E)^2 + \left(\frac{p}{A} - E\right)(Q-E) + \frac{EF}{2} = -W_3(p,Q,t) \tag{5c}$$

$$V_4 = -\frac{D}{2C}p^2 - \frac{A}{2C}(P-F)^2 + \left(\frac{p}{C} - E\right)(P-F) - \frac{EF}{2} = -W_4(p,P,t) \tag{5d}$$

These four functions $W_1 \sim W_4$ are sometimes called eikonals [3]. Because of the relationship between optics and mechanics, we prefer to call them generating functions [8], and these generate the affine transformation (1) by differentiation with respect to the canonical variables as follows,

$$W_1(q,Q,t) \quad \text{for } p = +\frac{\partial W_1}{\partial q}, P = -\frac{\partial W_1}{\partial Q}, \tag{6a}$$

$$W_2(q,P,t) \quad \text{for } p = +\frac{\partial W_2}{\partial q}, Q = +\frac{\partial W_2}{\partial P}, \tag{6b}$$

$$W_3(p,Q,t) \quad \text{for } q = -\frac{\partial W_3}{\partial p}, P = -\frac{\partial W_3}{\partial Q}, \tag{6c}$$

$$W_4(p,P,t) \quad \text{for } q = -\frac{\partial W_4}{\partial p}, Q = +\frac{\partial W_4}{\partial P}, \tag{6d}$$

We listed four types of the generating functions in (5). From the theoretical and experimental points of view, it sometimes happens that we cannot describe the affine transformation via one of them. For example, the affine transformation in (20) below has zero component in $C = 0$. In that case, we cannot use (5d), but the other ones are available. The relationship between these eikonals (generating functions) is depicted in Figure 1. The functions at the ends of the arrows are related to each other by a Legendre transformation. For example, from the relation (6a), we obtain the variable $Q$ in terms of $q$ and $P$,

$$P = -\frac{\partial W_1}{\partial Q} = \frac{D}{B}(Q-E) - \frac{q}{B} + F\,.$$

Substituting this relation into (5a) and $W_2 = W_1 + PQ$, we obtain (5b).

## 3. Kernel of the Affine Transformation

In this section, we consider the quantum mechanical version of the affine transformation.

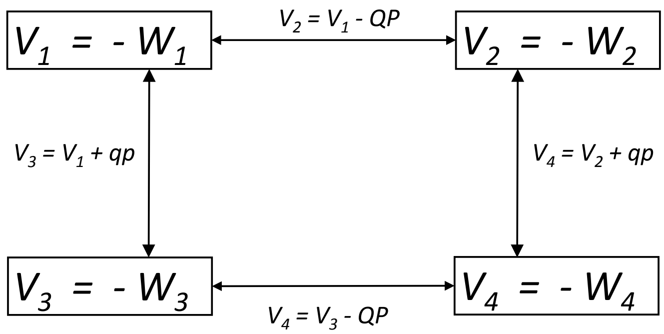



**Figure 1.** The eikonals (generating functions) are disposed on the corners of the square. The functions at the ends of arrows are related to each other by a Legendre transformation.

Corresponding to the canonical transformation in classical mechanics, the unitary transformation plays a central role in quantum mechanics. Analogous to the classical affine transformation (1), we define the quantum mechanical affine transformation as follows,

$$\hat{Q} = \hat{T}^\dagger \hat{q} \hat{T} = A\hat{q} + B\hat{p} + E , \tag{7a}$$

$$\hat{P} = \hat{T}^\dagger \hat{p} \hat{T} = C\hat{q} + D\hat{p} + F , \tag{7b}$$

where $\hat{}$ describes the q-number and $\hat{T}$ is a unitary operator which generates the affine transformation. Here, $AD - BC = 1$ is also needed when the canonical commutation relations $[\hat{q}, \hat{p}] = [\hat{Q}, \hat{P}] = i$ are satisfied.

To obtain the unitary operator $\hat{T}$, we introduce two operators. One is the displacement operator $\hat{D}$,

$$\hat{D} = \exp[-iE\hat{p} + iF\hat{q}] \tag{8}$$

which generates the displacements in position and momentum,

$$\hat{D}^\dagger \hat{q} \hat{D} = \hat{q} + E , \tag{9a}$$

$$\hat{D}^\dagger \hat{p} \hat{D} = \hat{p} + F . \tag{9b}$$

The other one is the unitary transformation $\hat{U}$,

$$\hat{U} = \exp\left[-i\left\{\alpha\frac{\hat{p}^2}{2} + \beta\frac{\hat{q}\hat{p} + \hat{p}\hat{q}}{2} + \gamma\frac{\hat{q}^2}{2}\right\}\right], \tag{10}$$

which generates the *ABCD*-transformation [9]

$$\hat{U}^\dagger \hat{q} \hat{U} = A\hat{q} + B\hat{p} , \tag{11a}$$

$$\hat{U}^\dagger \hat{p} U = C\hat{p} + D\hat{p} , \tag{11b}$$

where

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \cosh\Delta + \dfrac{\beta}{\Delta}\sinh\Delta & \dfrac{\alpha}{\Delta}\sinh\Delta \\ -\dfrac{\gamma}{\Delta}\sinh\Delta & \cosh\Delta - \dfrac{\beta}{\Delta}\sinh\Delta \end{pmatrix}, \tag{12}$$

and $\Delta^2 = \beta^2 - \alpha\gamma$. Note here that when we assign

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \cosh\xi + \cos\theta\sinh\xi & \sin\theta\sinh\xi \\ \sin\theta\sinh\xi & \cosh\xi - \cos\theta\sinh\xi \end{pmatrix}, \tag{13}$$

then $\hat{U}$ describes the squeezed operator [10].

We consider the unitary operator $\hat{T}$, which generates the quantum affine transformation (7), as a product of $\hat{D}$ and $\hat{U}$,

$$\hat{T} = \hat{D}\hat{U}. \tag{14}$$

Indeed, we obtain

$$\hat{Q} = \hat{T}^\dagger \hat{q} \hat{T} = \hat{U}^\dagger \hat{D}^\dagger \hat{q} \hat{D}\hat{U} = \hat{U}^\dagger (\hat{q} + E)\hat{U} = A\hat{q} + B\hat{p} + E, \tag{15a}$$

$$\hat{P} = \hat{T}^\dagger \hat{p} \hat{T} = \hat{U}^\dagger \hat{D}^\dagger \hat{p} \hat{D}\hat{U} = \hat{U}^\dagger (\hat{p} + F)\hat{U} = C\hat{q} + D\hat{p} + F, \tag{15b}$$

which is the quantum mechanical affine transformation (7) as we expected.

Now, we calculate the kernel of the affine transformation. The kernel is just the transition amplitude from the position $q$ at an initial time to the position $Q$ at a later time given by $\mathcal{K}^{affine}(Q,q) = \langle Q|\hat{T}|q\rangle$. To obtain the kernel, we use the coordinate identity operator $\int \mathrm{d}q' |q'\rangle\langle q'| = 1$ between the displacement operator $\hat{D}$ and the unitary operator $\hat{U}$:

$$\langle Q|\hat{T}|q\rangle = \langle Q|\hat{D}\hat{U}|q\rangle = \int \mathrm{d}q' \langle Q|\hat{D}|q'\rangle\langle q'|\hat{U}|q\rangle \tag{16}$$

Using the formulae

$$\hat{D}|q\rangle = |q+E\rangle \exp\left[+iFq + \frac{i}{2}EF\right], \tag{17a}$$

$$\hat{D}|p\rangle = |p+F\rangle \exp\left[-iEp - \frac{i}{2}EF\right], \tag{17b}$$

we obtain

$$\begin{aligned} \langle Q|\hat{T}|q\rangle &= \int \mathrm{d}q' \langle Q|q'+E\rangle\langle q'|\hat{U}|q\rangle \exp\left[iFq' + \frac{i}{2}EF\right] \\ &= \langle Q-E|\hat{U}|q\rangle \exp\left[iF(Q-E) + \frac{i}{2}EF\right]. \end{aligned} \tag{18}$$

Substituting the transition amplitude in terms of $\hat{U}$ [9], we obtain the result:

$$\langle Q|\hat{T}|q\rangle = \sqrt{\frac{1}{2\pi i B}} \exp\left[-i\left\{-\frac{A}{2B}q^2 - \frac{D}{2B}(Q-E)^2 + \left(\frac{q}{B} - F\right)(Q-E) - \frac{EF}{2}\right\}\right]. \tag{19}$$

We include the "irrelevant" constant phase factor which has often been neglected in the literature [11] [12]. The function in the exponent is in the same form as that of the generating function (5a). For example, let us consider a particle with mass $m$, subjected to a constant external force $f$, moving from $(q,p)$ to $(Q,P)$ in time $t$. The exact solution for this problem is described in the following,

$$\begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} 1 & \frac{t}{m} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} + \begin{pmatrix} \frac{ft^2}{2m} \\ ft \end{pmatrix}. \tag{20}$$

Substituting these parameters into (19), we obtain

$$\langle Q|\hat{T}|q\rangle = \sqrt{\frac{m}{2\pi it}}\exp\left[-i\left\{-\frac{m}{2t}(q-Q)^2 - \frac{ft}{2}(q+Q) + \frac{f^2t^3}{8m}\right\}\right], \tag{21}$$

which is the same equation as that obtained from the path integral [13].

The other kernels are derived in the same manner. We list all four types of transition amplitude below:

$$\langle Q|\hat{T}|q\rangle = \sqrt{\frac{1}{2\pi iB}}\exp\left[-iW_1(q,Q)\right] = \sqrt{\frac{1}{2\pi iB}}\exp\left[iV_1\right], \tag{22a}$$

$$\langle P|\hat{T}|q\rangle = \sqrt{\frac{1}{2\pi D}}\exp\left[-iW_2(q,P)\right] = \sqrt{\frac{1}{2\pi D}}\exp\left[iV_2\right], \tag{22b}$$

$$\langle Q|\hat{T}|p\rangle = \sqrt{\frac{1}{2\pi A}}\exp\left[-iW_3(p,Q)\right] = \sqrt{\frac{1}{2\pi A}}\exp\left[iV_3\right], \tag{22c}$$

$$\langle P|\hat{T}|p\rangle = \sqrt{\frac{-1}{2\pi iC}}\exp\left[-iW_4(p,P)\right] = \sqrt{\frac{-1}{2\pi iC}}\exp\left[iV_4\right], \tag{22d}$$

where the $W$'s in the exponentials are the generating functions (5) which generate the canonical transformation (1).

It is worth commenting here that it is well known in classical mechanics [8] that the generating functions (5) are related to each other by a Legendre transformation (**Figure 1**), whereas the kernels (22) are related to each other by a Fourier transformation. These relations are depicted in **Figure 2**.

## 4. Wavization of Gabor Function

Quantum mechanics is obtained by the "quantization" of classical mechanics. Similarly, physical optics is constructed by the "wavization" of geometrical optics [3] [6]. The famous example is that of Fraunhofer diffraction obtained by wavization of a plain wave. Let us consider the Gabor function [6] [14];

$$\langle q|Gabor\rangle = \left(\frac{1}{\pi\sigma^2}\right)^{1/4}\exp\left[ip_0\left(q-\frac{q_0}{2}\right)-\frac{1}{2\sigma^2}(q-q_0)^2\right], \tag{23}$$

where $p_0$ is the wave number and $\langle\hat{q}\rangle = q_0$ is the center of this wave packet. The

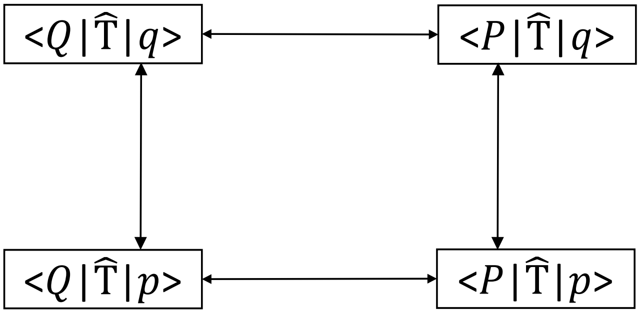



**Figure 2.** The kernels are disposed on the corners of the square. The functions at the ends of arrows are related to each other by a Fourier transformation.

width is obtained from

$$\Delta q = \sqrt{\langle \hat{q}^2 \rangle - \langle \hat{q} \rangle^2} = \frac{\sigma}{\sqrt{2}}. \tag{24}$$

To make the calculation easier, this wave packet (23) can be rewritten in the form,

$$\langle q | Gabor \rangle = \left( \frac{1}{\pi\sigma^2} \right)^{1/4} \exp\left[ -\frac{q^2}{2\sigma^2} + \frac{\sqrt{2}z}{\sigma} q - \frac{z^2}{2} - \frac{|z|^2}{2} \right], \tag{25}$$

with $z = \frac{1}{\sqrt{2}}\left( \frac{q_0}{\sigma} + i\sigma p_0 \right)$. Note that when $\sigma = 1$, it gives the position-representation of the coherent state wave function $\langle q | z \rangle$. It is also worth writing down the Fourier transformation of (23) and (25),

$$\langle p | Gabor \rangle = \left( \frac{\sigma^2}{\pi} \right)^{1/4} \exp\left[ -iq_0 \left( p - \frac{p_0}{2} \right) - \frac{\sigma^2}{2} (p - p_0)^2 \right], \tag{26}$$

$$= \left( \frac{\sigma^2}{\pi} \right)^{1/4} \exp\left[ -\frac{\sigma^2}{2} p^2 - i\sqrt{2}\sigma z p + \frac{z^2}{2} - \frac{|z|^2}{2} \right]. \tag{27}$$

Using this expression, we obtain $\langle \hat{p} \rangle = p_0$ and the width

$$\Delta p = \sqrt{\langle \hat{p}^2 \rangle - \langle \hat{p} \rangle^2} = \frac{1}{\sqrt{2}\sigma}. \tag{28}$$

This result with (24) gives

$$\Delta q \cdot \Delta p = \frac{1}{2}, \tag{29}$$

that is, the Gabor function satisfies the minimum uncertainty relation.

We obtain the affine transformation of the Gabor wave packet by using the kernel (22),

$$\begin{aligned} \langle Q | \hat{T} | Gabor \rangle &= \int \mathrm{d}q \langle Q | \hat{T} | q \rangle \langle q | Gabor \rangle \\ &= \left( \frac{1}{\pi\sigma^2 G^2} \right)^{1/4} \exp\left[ -\frac{H}{2\sigma^2 G} (Q - E)^2 + \left( \frac{\sqrt{2}z}{\sigma G} + iF \right) (Q - E) - \frac{G^*}{G} \frac{z^2}{2} - \frac{|z|^2}{2} + i\frac{EF}{2} \right], \end{aligned} \tag{30}$$

and

$$\begin{aligned} \langle P | \hat{T} | Gabor \rangle &= \int \mathrm{d}q \langle P | \hat{T} | q \rangle \langle q | Gabor \rangle = \int \mathrm{d}p \langle P | \hat{T} | p \rangle \langle p | Gabor \rangle \\ &= \left( \frac{\sigma^2}{\pi H^2} \right)^{1/4} \exp\left[ -\frac{\sigma^2 G}{2H} (P - F)^2 - i\left( \frac{\sqrt{2}\sigma z}{H} + E \right) (P - F) + \frac{H^*}{H} \frac{z^2}{2} - \frac{|z|^2}{2} - i\frac{EF}{2} \right], \end{aligned} \tag{31}$$

where we introduce two complex variables,

$$G = A + i\frac{B}{\sigma^2}, \tag{32a}$$

$$H = D - i\sigma^2 C. \tag{32b}$$

Having the probability density from (30) and (31), we obtain

$$\left|\langle Q|\hat{T}|Gabor\rangle\right|^2 = \sqrt{\frac{1}{\pi\sigma^2|G|^2}}\exp\left[-\frac{1}{\sigma^2|G|^2}\left\{Q-\left(Aq_0+Bp_0+E\right)\right\}^2\right], \tag{33a}$$

$$\left|\langle P|\hat{T}|Gabor\rangle\right|^2 = \sqrt{\frac{\sigma^2}{\pi|H|^2}}\exp\left[-\frac{\sigma^2}{|H|^2}\left\{P-\left(Cq_0+Dp_0+F\right)\right\}^2\right]. \tag{33b}$$

The center of the Gabor function propagates along the affine transformation (1). From these Equation (33), we obtain the variances

$$\Delta Q = \sqrt{\langle\hat{Q}^2\rangle-\langle\hat{Q}\rangle^2} = \frac{\sigma}{\sqrt{2}}|G|, \tag{34a}$$

$$\Delta P = \sqrt{\langle\hat{P}^2\rangle-\langle\hat{P}\rangle^2} = \frac{|H|}{\sqrt{2}\sigma}, \tag{34b}$$

and the uncertainty relation

$$\Delta Q\cdot\Delta P = \frac{1}{2}|G|\cdot|H| = \frac{1}{2}\sqrt{\left(A^2+\frac{B^2}{\sigma^4}\right)\left(D^2+\sigma^4C^2\right)}. \tag{35}$$

Since the only constraint for the parameters $(A,B,C,D)$ is $AD-BC=1$, these parameters have time dependency. So, these results (34) and (35), show the time development of the variances and the uncertainty relation of the Gabor function.

Let us show two examples here. As we saw in (25), the Gabor function with $\sigma=1$ signifies a coherent state. So, using (13) as $(A,B,C,D)$, (35) becomes

$$\Delta Q\cdot\Delta P = \frac{1}{2}\sqrt{1+\left(\sinh 2\xi\sin\theta\right)^2}, \tag{36}$$

which coincides with the uncertainty relation of the squeezed state [10]. The other example is where we use (20), then (35) becomes

$$\Delta Q\cdot\Delta P = \frac{1}{2}\sqrt{1+\frac{t^2}{\sigma^4m^2}}, \tag{37}$$

which coincides with the uncertainty relation [15] of the spreading of the Gaussian wave packet in time.

## 5. Affine Transformation of the Wigner Function

The Wigner function [16]-[18] is widely used in studying optics and the correspondence between classical and quantum mechanics [6] [7]. The Wigner function for any wave function $|\psi\rangle$ is defined by

$$f(Q,P) = \int \mathrm{d}u\,\mathrm{e}^{iuP}\left\langle Q-\frac{u}{2}\middle|\psi\right\rangle\left\langle\psi\middle|Q+\frac{u}{2}\right\rangle. \tag{38}$$

When we take a Gabor function (23) for any wave function $|\psi\rangle=|Gabor\rangle$, the Wigner function (38) becomes

$$f_G(Q,P) = 2\exp\left[-\frac{1}{\sigma^2}\left(Q-q_0\right)^2-\sigma^2\left(P-p_0\right)^2\right]. \tag{39}$$

Now we apply the unitary operator $\hat{T}$ to any wave function; $|\psi\rangle \to \hat{T}|\psi\rangle$. We obtain the affine transformation of the Wigner function,

$$f^{affine}(Q,P) = \int \mathrm{d}u\, \mathrm{e}^{iuP} \left\langle Q - \frac{u}{2} \middle| \hat{T} \middle| \psi \right\rangle \left\langle \psi \middle| \hat{T}^{\dagger} \middle| Q + \frac{u}{2} \right\rangle. \tag{40}$$

To cast the right hand side, we use the coordinate identity operator $\int \mathrm{d}x |x\rangle\langle x| = 1$ twice,

$$\int \mathrm{d}u \mathrm{d}x \mathrm{d}x' \mathrm{e}^{iuP} \left\langle Q - \frac{u}{2} \middle| \hat{T} \middle| x \right\rangle \langle x|\psi\rangle\langle\psi|x'\rangle \left\langle x' \middle| \hat{T}^{\dagger} \middle| Q + \frac{u}{2} \right\rangle. \tag{41}$$

Substituting the kernel (22a) into (41) and integrating over $u$, we obtain

$$\begin{aligned}
&\int \frac{\mathrm{d}x \mathrm{d}x'}{2\pi B} \langle x|\psi\rangle\langle\psi|x'\rangle \exp\left[ i\frac{A}{2B}\left(x^2 - x'^2\right) - i\frac{Q-E}{B}\left(x - x'\right) \right] \\
&\times 2\pi\delta\left( P - F - \frac{D(Q-E)}{B} + \frac{x+x'}{2B} \right),
\end{aligned} \tag{42}$$

where $\delta(z)$ is a delta function of $z$. Changing the variables $u = x - x'$ and $v = \frac{x+x'}{2}$, we obtain

$$\begin{aligned}
&\int \frac{\mathrm{d}u \mathrm{d}v}{2\pi B} \left\langle v + \frac{u}{2} \middle| \psi \right\rangle \left\langle \psi \middle| v - \frac{u}{2} \right\rangle \exp\left[ i\frac{A}{B}uv - i\frac{Q-E}{B}u \right] \\
&\times 2\pi B\delta\left( B(P-F) - D(Q-E) + v \right),
\end{aligned} \tag{43}$$

where we use the formula $\delta(ax) = a^{-1}\delta(x)$. Integrating over $v$, we obtain

$$\begin{aligned}
&f^{affine}(Q,P) = \int \mathrm{d}u\, \mathrm{e}^{iu\{A(P-F)-C(Q-E)\}} \\
&\times \left\langle D(Q-E) - B(P-F) - \frac{u}{2} \middle| \psi \right\rangle \left\langle \psi \middle| D(Q-E) - B(P-F) + \frac{u}{2} \right\rangle,
\end{aligned} \tag{44}$$

where we use $AD - BC = 1$ and change the variable $u \to -u$. This is the affine transformation of the Wigner function which is a generalization of (38) and can be applied to any wave function $|\psi\rangle$ and to any affine transformation with the condition $AD - BC = 1$. Equation (44) shows that the $ABCD$-part describes the area-preserving distortion, and the $E$, $F$-part describes the displacement in phase space. It is permissible for any affine transformation to have time dependency, so we are able to investigate the dynamic properties of the Wigner function in phase space.

As an example of a wave function $|\psi\rangle$, we take a Gabor function. Substituting the Gabor function (23) into (44) and integrating over $u$, we obtain

$$\begin{aligned}
&f_G^{affine}(Q,P) \\
&= 2\exp\left[ -\frac{1}{\sigma^2}\left\{ D(Q-E) - B(P-F) - q_0 \right\}^2 - \sigma^2\left\{ A(P-F) - C(Q-E) - p_0 \right\}^2 \right].
\end{aligned} \tag{45}$$

This equation is a generalization of (39), that is, in its initial state, the Wigner function of the Gabor function is represented by (39). Once the affine transformation switches on, the Wigner function changes along with (45). Note that integrating (45) over $P$ and $Q$ respectively, we recover (33);

$$\int \frac{\mathrm{d}P}{2\pi} f_G^{affine}(Q,P) = \left|\langle Q|\hat{T}|Gabor\rangle\right|^2, \tag{46a}$$

$$\int \frac{\mathrm{d}Q}{2\pi} f_G^{affine}(Q,P) = \left|\langle P|\hat{T}|Gabor\rangle\right|^2, \tag{46b}$$

which is the correct character of the Wigner function.

## 6. Summary

We have developed the mathematical properties of an affine transformation from the optical and mechanical points of view. The kernels of the affine transformation were clearly derived and comprise the eikonals (generating functions) which generated the affine transformation in optics (mechanics).

Using the kernel, we discussed the wavization of the Gabor function. The Gabor function has a Gaussian profile and is symmetric in position and momentum. We found the time development of the uncertainty relation, according to the affine transformation.

We also discussed the affine transformation of the Wigner function and showed not only the distortion but also the dynamic movement of the Wigner function in phase space.

## References

[1] Hecht, E. and Zajac, A. (1974) Optics. Addison-Wesley, Boston.

[2] Mickelson, A.R. (1992) Physical Optics. Van Nostrand Reinhold, New York. http://dx.doi.org/10.1007/978-1-4615-3530-0

[3] Guillemin, V. and Sternberg, S. (1984) Symplectic Technique in Physics. Cambridge.

[4] Moshinsky, M. and Quesne, C. (1971) *Journal of Mathematical Physics*, **12**, 1772. http://dx.doi.org/10.1063/1.1665805

[5] Boon, M.H. and Seligman, T.H. (1973) *Journal of Mathematical Physics*, **14**, 1224. http://dx.doi.org/10.1063/1.1666470

[6] Torre, A. (2005) Linear Ray and Wave Optics in Phase Space. Elsevier, Amsterdam.

[7] Ozaktas, H.M., Zalevsky, Z. and Kutay, M.A. (2001) The Fractional Fourier Transformation. John Wiley & Sons, Hoboken.

[8] Goldstein, H. (1950) Classical Mechanics. Addison-Wesley, Boston.

[9] Ogura, A. (2009) *Journal of Physics*, **B42**, 145504. http://dx.doi.org/10.1088/0953-4075/42/14/145504

[10] Stoler, D. (1970) *Physical Review D*, **1**, 3217. http://dx.doi.org/10.1103/PhysRevD.1.3217

[11] Abe, S. and Sheridan, J.T. (1994) *Journal of Physics*, **A27**, 4179. http://dx.doi.org/10.1088/0305-4470/27/12/023

[12] Cai, L.Z. (2000) *Optics Communications* **185**, 271-276. http://dx.doi.org/10.1016/S0030-4018(00)01005-1

[13] Feynman, R.P. and Hibbs, A.R. (1965) Quantum Mechanics and Path Integral. McGraw-Hill Inc., Boston.

[14] Gabor, D. (1946) *J. of the IEE*, **93**, 429.

[15] Shiff, L.I. (1955) Quantum Mechanics. McGraw-Hill Inc., Boston.

[16] Wigner, E.P. (1932) *Physical Review*, **40**, 749. http://dx.doi.org/10.1103/PhysRev.40.749

[17] Kim, Y.S. and Wigner, E.P. (1990) *American Journal of Physics*, **58**, 439. http://dx.doi.org/10.1119/1.16475

[18] Hillery, M., O'connell, R.F., Scully, M.O. and Wigner, E.P. (1984) *Physics Reports*, **106**, 121-167. http://dx.doi.org/10.1016/0370-1573(84)90160-1

Journal of Modern Physics, 2016, 7, 1749-1761
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# LIGO Experiments Cannot Detect Gravitational Waves by Using Laser Michelson Interferometers

## —Light's Wavelength and Speed Change Simultaneously When Gravitational Waves Exist Which Make the Detections of Gravitational Waves Impossible for LIGO Experiments

**Xiaochun Mei[1], Zhixun Huang[2], Policarpo Yōshin Ulianov[3], Ping Yu[4]**

[1]Institute of Innovative Physics in Fuzhou, Fuzhou, China
[2]Communication University of China, Beijing, China
[3]Equalix Tecnologia LTDA, Florianópolis, Brazil
[4]Cognitech Calculating Technology Institute, Los Angeles, CA, USA
Email: ycwlyjs@yeah.net, huangzhixun75@163.com, policarpoyu@gmail.com, yupingpingyu@yahoo.com

**How to cite this paper:** Mei, X.C., Huang, Z.X., Ulianov, P.Y. and Yu, P. (2016) LIGO Experiments Cannot Detect Gravitational Waves by Using Laser Michelson Interferometers. *Journal of Modern Physics*, **7**, 1749-1761.
http://dx.doi.org/10.4236/jmp.2016.713157

**Received:** July 28, 2016
**Accepted:** September 25, 2016
**Published:** September 28, 2016



## Abstract

It is proved strictly based on general relativity that two important factors are neglected in LIGO experiments by using Michelson interferometers so that fatal mistakes were caused. One is that the gravitational wave changes the wavelength of light. Another is that light's speed is not a constant when gravitational waves exist. According to general relativity, gravitational wave affects spatial distance, so it also affects the wavelength of light synchronously. By considering this fact, the phase differences of lasers were invariable when gravitational waves passed through Michelson interferometers. In addition, when gravitational waves exist, the spatial part of metric changes but the time part of metric is unchanged. In this way, light's speed is not a constant. When the calculation method of time difference is used in LIGO experiments, the phase shift of interference fringes is still zero. So the design principle of LIGO experiment is wrong. It was impossible for LIGO to detect gravitational wave by using Michelson interferometers. Because light's speed is not a constant, the signals of LIGO experiments become mismatching. It means that these signals are noises actually, caused by occasional reasons, no gravitational waves are detected really. In fact, in the history of physics, Michelson and Morley tried to find the absolute motion of the earth by using Michelson interferometers but failed at last. The basic principle of LIGO experiment is the same as that of Michelson-Morley experiment in which the phases of lights were invariable. Only zero result can be obtained, so LIGO experiments are destined failed to find gravitational waves.

## Keywords

Gravitational Wave, LIGO Experiment, General Relativity, Special Relativity, Michelson Interferometer, Michelson-Morley Experiment, GW150914, WG151226

## 1. Introduction

February 11, 2016, LIGO (Laser Interference Gravitational-Waves Observatory) announced to detect gravitational waves events GW150914 [1]. Four months later, they announced to detect another two gravitational events WG151226 and LVT151012 [2]. In LIGO experiments, Michelson laser interferometers were used. Based on general relativity, we proved strictly that by using Michelson interferometers, LIGO cannot detect gravitational waves. The basic principle of LIOGO experiment is wrong. The so-called detections of gravitational waves and the observations of binary black hole mergers are impossible.

The design principle of LIGO experiments is as follows. According to general relativity, gravitational waves stretch and compress space to change the lengths of interferometer's arms. When two lights travelling along two arms which are displaced vertically meet together, the shapes of interference fringes will change. Based on this phase shifts, gravitational waves can be observed.

There are two methods to calculate the phase shift of interference fringes in classical optics. One is to calculate the phase difference of two lights and another is to calculate the time difference of two lights when they arrive at the screen. In LIGO experiments, two of them were used. But the calculations are based on a precondition, *i.e.*, light's speed is a constant.

As well-known, light's phase is related to its wavelength. The stretch and squeeze of space also cause the change of light's wavelength and affect phases. However, LOGO experiment neglected the effect of gravitational wave on the wavelength of light. If the effects of gravitational wave on light's wavelength and interferometer arm's lengths are considered simultaneously, light's phases are unchanged in Michelson interferometers. So it is impossible for LIGO experiments to detect gravitational waves.

On the other hand, light's speed was considered as a constant in LIGO experiments. It is proved strictly based on general relativity that when gravitational waves exist, light's speed is not a constant again. If light's speed is less than its speed in vacuum when it travels along one arm of interferometer, its speed will be great than its speed in vacuum when it travels along another arm, *i.e.*, so-called superluminal motion occurs. In this way, no time differences exist when two lights meet together in Michelson interferometer. Therefore, according to the second method of calculation, LIGO experiments did not detect gravitational waves too.

The other principle problems existing in LIGO experiments are briefly discussed in this paper. The conclusion is that LIGO experiments do not detect gravitational waves and no binary black hole mergers are observed. The signals occurred in LIGO experiments could only be noises caused by some occasional reasons.

## 2. Light's Phase Difference Is Invariable in LIGO Experiments

According to general relativity, under the condition of weak field, the metric tensor is

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\eta\nu}(x) \tag{1}$$

Here $\eta_{\mu\nu}$ is the metric of flat space-time and $h_{\mu\nu}(x)$ is a small quantity. Substitute (1) in the Einstein's equation of gravitational field, it can be proved that the modal of gravitational radiation is quadrupole moment. In a small region, we may assume $h_{\eta\nu}(x) = h_{\eta\nu}(t)$. When gravitational wave propagates along the *x*-axis, the intensity of gravitational field is $h_{11}(t)$. While it propagates along the *y*-axis, the intensity is $h_{22}(t)$. It can be proved to have relation $h_{11}(t) = -h_{22}(t)$ [3].

On the other hand, according to general relativity, we have $\mathrm{d}s^2 = 0$ for light's motion. Suppose that gravitational wave propagates along the *z*-axis, when lights propagate along the *x*-axis and the *y*-axis individually, we have [4]

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \left[1 + h_{11}(t)\right]\mathrm{d}x^2 = 0 \tag{2}$$

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \left[1 + h_{22}(t)\right]\mathrm{d}y^2 = 0 \tag{3}$$

It is obvious that time is flat but space is curved according (2) and (3). The propagation forms of light are changed when gravitational waves exist. Due to $|h_{11}| \ll 1$, $|h_{22}| \ll 1$ and $h_{11}(t) = -h_{22}(t)$, we have

$$\mathrm{d}x = \frac{c}{\sqrt{1+h_{11}}}\mathrm{d}t = c\left(1 - \frac{1}{2}h_{11}(t)\right)\mathrm{d}t \tag{4}$$

$$\mathrm{d}y = \frac{c}{\sqrt{1+h_{22}}}\mathrm{d}t = c\left(1 + \frac{1}{2}h_{11}(t)\right)\mathrm{d}t \tag{5}$$

LIGO experiments used Michelson interferometers to detect gravitational waves. The principle of Michelson interferometer is shown in Figure 1. Light is emitted from the source S and split into two beams by beam splitter O. Light 1 passes through O, arrives at reflector $M_1$ and is reflected by $M_1$ and O, then arrived at *E*. Light 2 is reflected by

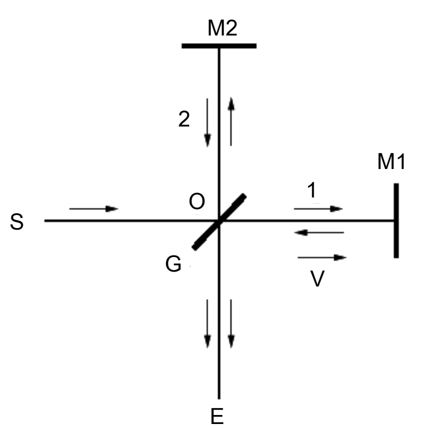



Figure 1. The principle of Michelson interferometers.

O, arrives at $M_2$ and is reflected, then arrived at $E$ too. Two lights overlay and form interference fringes which can be observed by observer at $E$.

In order to reveal the problems of LIGO experiments clearly, we discuss the simplest situation. Suppose that the length of interferometer's arm is $L_0$ and let $h_{11}(t) = h =$ constant. The time interval is $t_2 - t_1 = 2\tau$ when the light moves a round-trip along the arm. The integral of (4) and (5) are

$$x = 2L_0\left(1 - h/2\right),\ \ y = 2L_0\left(1 + h/2\right) \tag{6}$$

Here $L_0 = c\tau$. So the optical path difference is $\Delta L = y - x = 2L_0 h$ for lights move along two interferometer's arms. Suppose that the electric fields of lasers are

$$E_x = E_0\cos\left(\omega t - kx\right),\ \ E_y = E_0\cos\left(\omega t - ky\right) \tag{7}$$

Here, $k = 2\pi/\lambda$, $\omega = 2\pi\nu$ and $\nu\lambda = c$. According to classical optics, by adding two amplitudes together directly and taking square, we obtain light's intensity which is unrelated to time

$$E^2 = \left(E_x + E_y\right)^2 = 2E_0^2\left(1 + \cos\Delta\delta\right) \tag{8}$$

The difference of phases is

$$\Delta\delta = k\left(y - x\right) = \frac{2\pi}{\lambda}\left(y - x\right) \tag{9}$$

If there is no gravitational wave, we have $y = x = 2L_0$ and get $\Delta\delta = 0$. If there is a gravitational wave which passes through the interferometers, according to the current theory, the difference of phases is

$$\Delta\delta = \frac{2\pi}{\lambda}\left(y - x\right) = \frac{4\pi L_0 h}{\lambda} \neq 0 \tag{10}$$

Therefore, gravitational waves would cause the phase changes of interference fringes. By observing the change, gravitational waves would be detected.

However, the calculation above has serious defects. At first, according to general relativity strictly, the formulas (1) and (2) are only suitable for two particles in vacuum without the existence of electromagnetic interaction. In LIGO experiments, two mirrors are hanged in interferometers using fiber material. Interferometers are fixed on the steal tubers which are fixed on the surface of the earth. Whole system is controlled by electromagnetic interaction. As we known, the intensity of electromagnetic interaction is $10^{40}$ times greater than gravitational interaction. Therefore, gravitational waves cannot overcome electromagnetic forces to change the length of interferometer's aims or make two mirrors vibration by overcoming the stain forces acted on fiber material. This is just the reason why J. Weber's gravitational wave experiments failed. This is the critical defect of LIGO experiments. We have discussed this problem in Document [5], so we do not discuss it any more here.

Second, the major point in this paper is to emphasize that the effect of gravitational wave on the wavelength of light has not been considered in LIGO experiments. In fact, if gravitational wave causes the change of spatial distance, it also causes the change of light's wavelength. Both are synchronous. According to (6), when gravitational waves exist, the wavelengths of lights should become

$$\lambda_x = \lambda(1 - h/2),\ \ \lambda_y = \lambda(1 + h/2) \tag{11}$$

when two lights meet together, the difference of phases should be

$$\Delta\delta = 2\pi\left(\frac{y}{\lambda_y} - \frac{x}{\lambda_x}\right) = 2\pi\left(\frac{2L_0}{\lambda} - \frac{2L_0}{\lambda}\right) = 0 \tag{12}$$

Therefore, interference fringes are unchanged. That is to say, it is impossible to detect gravitational waves by using Michelson interferometers. If $h_{11}(t) \neq \text{constant}$, we write it as

$$h_{11}(t) = h\sin(\Omega t + \theta_0) \tag{13}$$

Here, $\Omega$ is the frequency of gravitational wave. Substitute (13) in (5) and (6), the integrals become

$$\begin{aligned} x &= 2c\tau - \frac{ch}{2}\int_0^{2\tau} \sin(\Omega t + \theta_0)\mathrm{d}t \\ &= L_0 - \frac{ch}{2\Omega}\left[\cos(2\Omega\tau + \theta_0) - \cos\theta_0\right] = L_0(1 - A/2) \end{aligned} \tag{14}$$

$$\begin{aligned} y &= 2c\tau + \frac{ch}{2}\int_0^{2\tau} \sin(\Omega t + \theta_0)\mathrm{d}t \\ &= L_0 + \frac{ch}{2\Omega}\left[\cos(2\Omega\tau + \theta_0) - \cos\theta_0\right] = L_0(1 + A/2) \end{aligned} \tag{15}$$

Here

$$A = \frac{ch}{\Omega L_0}\left[\cos(2\Omega\tau + \theta_0) - \cos\theta_0\right] \tag{16}$$

The result is the same with (6) by substituting $A$ for $h$.

In LIGO experiments, by assuming that gravitational wave's speed is light's speed in vacuum, the frequency of gravitational wave is $\nu = 30 \sim 300\ \text{Hz}$ and the wavelength is $\lambda = c/\nu = 3\times10^6\ \text{m}$. The length of interferometer's arm is $L_0 = 4\times10^3\ \text{m}$, so we have $\lambda \gg L_0$. In the extent of interferometer size, the wavelength of gravitational wave can be considered as a fixed value. The formula (11) is still suitable by substituting $A$ for $h$. So, even though (13) was used to describe gravitational waves, LIGO experiments could not detect gravitational waves too.

## 3. Light's Speed Is Not a Constant When Gravitational Waves Exist

Based on (4) and (5), we can obtain an important conclusion, *i.e.*, light's speed is not a constant again when gravitational waves exist

$$\begin{aligned} V_x &= \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{c}{\sqrt{1 + h_{11}}} \approx c\left(1 - \frac{1}{2}h_{11}\right) \neq c \\ V_y &= \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{c}{\sqrt{1 + h_{22}}} = c\left(1 + \frac{1}{2}h_{11}\right) \neq c \end{aligned} \tag{17}$$

This result also causes a great effect on LIGO experiments. The current theory always considers light's speed to be a constant in gravitational fields. According to Reference [4], (17) means that the spatial refractive index becomes $1+h_{kk}/2$ from 1 due to the existence of gravitational waves. In this medium space, light's speed is changed. More interesting is that if $h_{11}>0$, we have $V_x<c$ and $V_y>c$. That is to say, $V_y$ exceeds light's speed in vacuum. How do we explain this result? No one consider this problem at present.

Reference [4] also indicates that "For Gaussian beam, the interval of space-time is not equal to zero. In the laser detectors of gravitational waves, Gaussian beams are used. Do these lights exist in curved space-time?" [4]. According to strict calculation, when gravitational waves exist, the propagation speed of Gaussian beam is

$$V_c = c\left|1-\frac{2}{(k\omega_0)^2+(2z/\omega_0)^2}\right|<c \tag{18}$$

Here, $\omega_0^2$ is the spot size of Gaussian beam, $k$ is the absolute value of wave's vector and $z$ is the coordinate of light beam. These results will cause great influence on the waveform match in LIGO experiment. The original matching signals would become mismatching when comparing them with the templates of waveforms. The conclusion to detect gravitational waves should be reconsidered.

In fact, LIGO team also admits that gravitational wave changes light's wavelength. In the LIGO's FAQ page (https://www.ligo.caltech.edu/page/faq) we can see the following question:

"*If a gravitational wave stretches the distance between the LIGO mirrors, doesn't it also stretch the wavelength of the laser light?*"

The answer of LIGO team is:

"*A gravitational wave does stretch and squeeze the wavelength of the light in the arms. But the interference pattern doesn't come about because of the difference between the length of the arm and the wavelength of the light. Instead it's caused by the different arrival time of the light wave's "crests and troughs" from one arm with the arrival time of the light that traveled in the other arm. To get how this works, it is also important to know that gravitational waves do NOT change the speed of light.*

The answer is very confusing, showing that they aware of the problem but try to escape from it. Then they say

"*But the interference pattern doesn't come about because of the difference between the length of the arm and the* wavelength *of the light.*"

The sentence makes no sense. In the above explanation, we see that

"*Instead it's caused by the different arrival time of the light wave's 'crests and troughs' from one arm with the arrival time of the light that traveled in the other* *arm. To get how this works, it is also important to know that gravitational waves do NOT change the speed of light.*"

In this sentence, LIGO emphases that gravitational waves do not change light's speed. This is the foundation of LIGO experiments. Because this conclusion does not hold, LIGO's explanation is untenable.

It is a confused problem for many physicists whether or not light's speed is a constant in gravitational field. To measurement speed, we first need to have unit ruler and unit clock. According to general relativity, gravitational field cause space-time curved. In the gravitational field, we have too definitions for ruler and clock, *i.e.*, coordinate ruler and coordinate clock, as well as standard ruler and standard clock (or proper ruler and clock). Coordinate ruler and coordinate clock are fixed at a certain point of gravitational field. They vary with the strength of gravitational field. Standard ruler and standard clock are fixed on the local reference frame which falls free in gravitational field. In local reference frame, gravitational force is canceled, so standard ruler and standard clock are unchanged.

It has been proved that if the metric tensor $g_{0i}$ which is related to time is not equal to zero, *i.e.*, $g_{0i} \neq 0$, no matter what ruler and clock are used, light's speed is not a constant. If $g_{0i} = 0$, by using coordinate ruler and coordinate clock, light's speed is not a constant. Using standard ruler and standard clock, light's speed becomes a constant. But in this case, the observer is also located at the reference frame which falls free in gravitational field [3].

In LIGO experiments, observers located at a gravitational field caused by gravitational wave, rather than falling free in gravitational field, so what they used were coordinate ruler and coordinate clock. Therefore, light's speed in LIGO experiments are not a constant. In fact, according to (2), the time part of metric is flat and the spatial part is curved, so the speed $V_x = \mathrm{d}x/\mathrm{d}t$ is not a constant certainly.

According to this definition, by using coordinate ruler and coordinate clock, light's speeds in gravitational fields are generally less its speed in vacuum. For example, light's speed is $V_r = \mathrm{d}r/\mathrm{d}t = c\left(1-\alpha/r\right) < c$ in the gravitational field of spherical symmetry according to Schwarzschild metric and $V_r = c\sqrt{1-kr^2}/R < 1$ at present moment with $R = 1$ in the gravitational field of cosmology according to the R-W metric. But in the early period time of cosmos with $R < 1$, light's speed is great than its speed in vacuum. So it is not strange that light's speed may be greater than its speed in vacuum at a certain direction if gravitational waves exist.

## 4. Phases Shifts Cannot Be Obtained by the Calculation Method of Time Difference in LIGO Experiments

In classical optics, the difference of time is also used to calculate the change of interference fringes. However, it has a precondition, *i.e.*, light's speed is a constant. LIGO experiments used time differences to calculate the changes of interference images [6]. Due to the fact that light's speed is not a constant when gravitational waves exist, we prove that it is impossible to use time difference to calculate the change of interference fringes.

Thought the lengths of interferometer's arms change, the speed of light also changes synchronously, so that the time that light travels along the arms is unchanged too.

Because light's frequency is $\omega = 2\pi\nu = 2\pi c/\lambda$, when gravitational wave exists, if light's speed is unchanged but the wavelength changes, the frequency $\omega$ will change. In this case, (7) should be written as

$$E_x = E_0 \cos\left(\omega_x t - k_x x\right),\ E_y = E_0 \cos\left(\omega_y t - k_y y\right) \tag{19}$$

when two lights are superposed, we cannot get (8). The result is related to time and becomes very complex. If light's speed is not a constant, according to (6) and (11), we have

$$\omega_x = 2\pi\nu_x = \frac{2\pi V_x}{\lambda_x} = \frac{2\pi c}{\lambda} = \omega,\ \ \omega_y = 2\pi\nu_y = \frac{2\pi V_y}{\lambda_y} = \frac{2\pi c}{\lambda} = \omega \tag{20}$$

In this case, light's frequency is invariable and the formula (8) is still tenable. So, when gravitational waves exist, we should think that light moves in medium. Light's frequency is unchanged but its speed and wavelength change. Only in this way, we can reach consistency in physics and logic. In fact, (20) is well-found in classical physics. As mentioned in [7], in a static medium, wave's speed changes but frequency does not change, so wavelength also changes.

We know from (7) that light's phase is determined by both factors $\omega t$ and $kx$. Here $kx = 2\pi x/\lambda$ is an invariable quantity according to discussion above. Because of $\omega = 2\pi\nu = 2\pi/T$ and $T$ is the period of light which changes with $t$ synchronously. We always have $\omega' = 2\pi\nu' = 2\pi t'/T' = 2\pi t/T = \omega$. Because gravitational waves do not affect time $t$, the phase $\omega t$ of light is also unchanged in LIGO experiments.

## 5. The Problems Existing in the Third Method to Calculate Phase Shifts of Light

There is a more complex method to calculate the phase shift of light for LIGO experiments by considering interaction between gravitational field and electromagnetic field, or by solving the Maxwell's equations in a curved space caused by gravitational wave [8]. This method also has many problems. We discuss them briefly below.

In this calculation, two arms of interferometers are located at the $x$-axis and the $y$-axis. If there is no gravitational wave, the vibration direction of electric field is along the $y$-axis for the light propagating along the $x$-axis (electromagnetic wave is transverse wave), we have

$$E_y^{(0)} = E_0\left[\mathrm{e}^{i(kx-\omega t)} - \mathrm{e}^{-i(kx-\omega t-2ka)}\right] = -F_{02}^{(0)} \tag{21}$$

Here, $a$ is the coordinate of reflect mirror, $F_{ik}^{(0)}$ is electromagnetic tensor. The form of magnetic field is the same, so we do not write it out here. When the light propagates along the $y$-axis, the vibration of electric field is along the direction of $x$-axis with

$$E_x^{(0)} = E_0\left[\mathrm{e}^{i(ky-\omega t)} + \mathrm{e}^{-i(ky-\omega t-2ka)}\right] = -F_{01}^{(0)} \tag{22}$$

Meanwhile, gravitational wave propagates along the $z$-axis with

$$h_{11} = -h_{22} = -A\cos\left(k_g z - \omega_g t\right) \tag{23}$$

when gravitational wave exists, electromagnetic tensors become

$$F_{\mu\nu} = F_{\mu\nu}^{(0)} + F_{\mu\nu}^{(1)} \tag{24}$$

Here, $F_{\mu\nu}^{(1)}$ is a small quantity of electromagnetic field induced by gravitational field. Substitute (24) in the equation of electromagnetic field in curved space-time, the equation $F_{\mu\nu}^{(1)}$ satisfied is [9]:

$$F_{\mu\nu,\rho}^{(1)}\eta^{\rho\nu} = h_{\mu}^{\nu,\rho} F_{\nu\rho,\rho}^{(0)} + h^{\nu\rho} F_{\mu\nu,\rho}^{(0)} + O\left(h^2\right) \tag{25}$$

$$F_{\mu\nu,\rho}^{(1)} = F_{\nu\rho,\mu}^{(1)} + F_{\rho\mu,\nu}^{(1)} = 0 \tag{26}$$

By solving (25) and (26), the concrete form of $F_{\mu\nu}^{(1)}$ can be obtained and the phase shifts caused by gravitational waves can be determined. The phase shifts along two arms are [8]

$$\delta\varphi_x = \frac{A}{2}\frac{\omega}{\omega_g}\sin\omega_g\tau, \quad \delta\varphi_y = -\frac{A}{2}\frac{\omega}{\omega_g}\sin\omega_g\tau \tag{27}$$

The total phase shift between two arms is $\delta\varphi = \delta\varphi_x - \delta\varphi_y$. However, by careful analysis, we find following problems in this calculation.

1) This method is also based on the precondition that light's speed is unchanged. As proved above, this is impossible.

2) Because the phases of lights are not affected by gravitational waves, the forms of (21) and (22) are invariable when gravitational waves exist. We have $F_{\mu\nu}^{(1)} = 0$ in (24), no phase shifts of lights can be obtained by this calculating method.

3) According to (21) and (22), the vibration directions of two lights propagating along the *x*-axis and the *y*-axis are vertical, so they cannot interfere to each other. How did gravitational waves make the shifts of interference fringes? This is another basic problem for this calculation method.

In addition, the phase differences $\delta\varphi_x$ and $\delta\varphi_x$ caused by gravitational waves cannot be obtained independently and simultaneously by solving Equations (25) and (26). The author of the paper admitted that "we solve these equations in a special orientation which does not correspond to an actual interferometer arm" [9]. So the paper introduced "a fictitious system which is composed of an electromagnetic wave propagating along the *z* axis, …is perturbed by a gravitational wave moves along the *y*-axis." It means that the calculation did not consider the light propagating along another arm of interferometers.

After simplified calculation, a coordinate transformation was used to transform the result to original problem. For the light propagating along the *x*-axis, the coordinate transformations are $t' = t$, $x' = y$, $y' = z$ and $z' = x$ (The coordinate reference frame rotates 90 degrees around the *x*-axis, then rotates 90 degrees around the *z*-axis again along the clockwise directions.) For the light propagating along the *y*-axis, the coordinate transformations are $t' = t$, $x' = x$, $y' = z$ and $z' = -y$ (The coordinate reference frame rotates 90 degrees around the *x*-axis along the anticlockwise direction.).

In this way, two problems are caused.

1) When a light propagates along one arm, the interaction between gravitational wave and electromagnetic field is different from that when two lights propagate along two arms, or the formulas (25) and (26) are different in two situations. So this simplified method cannot represent real experiment processes.

2) After coordinate transformation, the electric field of light originally propagating along the $x$-axis becomes [8]

$$E_x'^{(0)} = E_0 \left[ \mathrm{e}^{i(kz'-\omega t')} - \mathrm{e}^{-i(kz'+\omega t'-2ka)} \right] \tag{28}$$

The electric field of light originally propagating along the $z$-axis becomes

$$E_x'^{(0)} = E_0 \mathrm{e}^{2ika} \left[ \mathrm{e}^{i(kz'-\omega t')} - \mathrm{e}^{-i(kz'+\omega t'-2ka')} \right] \tag{29}$$

The gravitational waves become

$$h_{11} = -h_{33} = -A\cos\left(k_g y' - \omega_g t'\right) \tag{30}$$

It is obvious that though the vibration directions of two lights become the same so that the interference fringes can be created, two lights move along the same directions. The process is inconsistent with real experiments of Michelson interferometers. That is to say, it is hard for this calculating method to reach consistence.

In fact, the result of this calculation contracts with the calculation in this paper. The method of this paper is standard one with clear image and definite significance in physics. If the results are different from it by using other methods, we should consider whether or not other methods are correct.

It is obvious that there are so many foundational problems in theory of LIGO experiments. It is meaningless to declare the detection of gravitational waves. Even thought the experiments are moved to space in future, it is still impossible to detect gravitational wave if Michelson interferometers are used.

## 6. Comparison between LIGO Experiment and Michelson-Morley Experiment

The principle of detecting gravitational wave by using Michelson interferometers was first proposed by M. E. Gertenshtein and V. I. Pustoit in early 1960s [8] and G. E. Moss, etc. in 1970s [9]. However, before Einstein put forward special relativity, A. A. Michelson and E.W. Money spent decades to conduct experiments by using Michelson interferometer, trying to find the absolute movement of the earth but failed at last. This result led to the birth of Einstein's special relativity. The explanation of special relativity for this zero result is based on the length contraction of interferometer. When one arm which moved in speed $V$ contracted, another arm which was at rest was unchanged. The speed of light was considered invariable in the process so that no any shift of interference fringes was observed.

It is obvious the principle of LIGO experiment is the same as that in Michelson experiments. Because Michelson's experiments could not find the changes of interference fringes, it is destined for LIGO experiments impossible to find gravitational waves [10].

We discuss this problem in detail. Suppose that the interferometer's arm is located along the *y*-axis and the arm along the *x* axis moves in speed $V$. For an observer who is at rest with the *y*-axis, the length contraction and time delay of the arm along the *x*-axis are

$$x' = x\sqrt{1-V^2/c^2}, \quad t' = t\Big/\sqrt{1-V^2/c^2} \tag{31}$$

Suppose that the period is $T'$ and the frequency is $\nu'$ for a light moving along the *x*-axis, we have $\nu'T'=1$, $\omega' = 2\pi\nu' = 2\pi/T'$, as well as $T' = T\Big/\sqrt{1-V^2/c^2}$ (period is also time). So we have

$$\omega't' = \frac{2\pi t'}{T'} = \frac{2\pi t}{T} = \omega t, \ \ k'x' = \frac{2\pi x'}{\lambda'} = \frac{2\pi x}{\lambda} = kx \tag{32}$$

It means that in the rotation processes of Michelson interferometers, the phase $\omega t - kx$ of light is unchanged. In this way, the absolute movement of the earth cannot be observed. The key is that light's speed is unchanged, frequency and wavelength change simultaneously in the processes. But in LIGO experiment, as shown in (2), (3) and (17), due to the fact that the time part of metric is flat but space is curved, light's speed and wavelength had to change when gravitational waves exist. This is just the difference between LIGO experiments and Michelson experiments. But the phases of lights are invariable in both experiments. We can only obtain zero results, so LIGO experiments are destined failed to find gravitational waves.

Let's make further calculation. The speed that the earth moves around the sun is $V = 3\times10^4$ m/s. The length of Michelson interferometer's arm is $L = 10$ m. According to special relativity, the Lorentz contraction of one arm in Michelson experiments is

$$\Delta L = L\left(1-\sqrt{1-V^2/c^2}\right) = L\times V^2\Big/\left(2c^2\right) = 5\times10^{-8}\ \text{m} \tag{33}$$

In LIGO experiment, the length change of arm is about $10^{-18}$ m, about one 20 billionth times smaller than that in Michelson experiment. Suppose that the shift of interference fringes can be observed in Michelson experiments. According to classical mechanics, the number of fringe shifts is about 0.2. Suppose that IGO experiments can detect the shift of interference fringes caused by gravitational waves, the number of fringe shifts is only one 100 billionth of Michelson experiment. How could LIGO experiments separate such small shifts from strong background noises of environment (including temperature influence) and identified that they were really the effect of gravitational waves?

In fact, LIGO's interferometers are fixed on two huge steel tubes with length 4000 m. The steel tubes are fixed on the surface of the earth under wind and rain. It is impossible to put so huge interferometers in a constant temperature rooms. The tubes are displaced vertically and 4000 m is not an ignorable length. The differences of temperatures exist and change with time frequently. Suppose that at a certain moment, the temperature of one tube changes 0.001 degree in one second. This is a conservative estimation. We calculate its influence on LIGO's experiment.

The expansion coefficient of common steel tube is $1.2 \times 10^{-5}$ m/degree. When temperature changes 0.001 degree, the change of tube length is $1.2 \times 10^{-5} \times 0.001 \times 4000 \approx 5 \times 10^{-5}$ m in one second. The action time of gravitational wave is 1 second. In this time, the length change of tube caused by gravitational wave is $10^{-18}$ m. The length change of tube caused by gravitational wave only is $2 \times 10^{-12}$ times less than that caused by the change of temperature.

What is this concept? It means that LIGO used a ruler of 10 Km to measure the radius of an atom. The length changes caused by temperature completely cover up the length changes caused by gravitational waves. No any reaction can be found when a signal of gravitational wave hit the interferometers of LIGO. LIGO's instrument cannot separate the effect of gravitational waves from temperature's effect. The SNR (signal to noise ratio) of 13 and 24 declaimed by LIGO is only an imaginary value in theory, having nothing to do with practical measurements.

## 7. Conclusions

In this paper, based on general relativity, we strictly prove that the LIGO experiments neglect two factors. One is the effects of gravitational waves on the wavelengths of light. Another is that light's speed is not a constant when gravitational waves exist. If these factors are considered, no phase shifts or interference fringe's changes can be observed in LIGO experiments by using Michelson interferometers.

In fact, in the laser detectors of gravitational waves, Gaussian beams are commonly used. The propagation speed of Gaussian beam is not a constant too. So the match of signals becomes a big problem without considering these factors in LIGO experiments.

In addition, in Reference [5], X. Mei and P. Yu pointed out that no source of gravitational wave burst was found in LIGO experiments. The so-called detections of gravitational waves were only a kind of computer simulation and image matching. LIGO experiments had not verified general relativity. The argument of LIGO team to verify the Einstein's prediction of gravitational wave was a vicious circle and invalid in logic. The method of numerical relativity to calculate the binary black hole mergers was incredible because too many approximations were involved.

In Reference [11], P. Ulianov indicated that the signals appeared in LIGO experiments may be caused by the changes of frequency in the US power grid. The analysis shows that one of noise sources in LIGO's detectors (32.5 Hz noise source) is connected to the 60 Hz power grid and at GW150914 event. This noise source presents an unusual level change. Besides that, the 32.5 Hz noise waveform is very similar to the gravitational waveform, found in GW150914 event. As LIGO system only monitored the power grid voltage levels without monitoring the 60 Hz frequency changes, this kind of changes over US power grid (that can affect both LIGO's detectors in a same time windows) was not perceived by the LIGO team.

Based on the arguments above, we can conclude that it is impossible for LIGO to detect gravitational waves. What they found may be some noises by some occasional reasons. So called finding of gravitational waves is actually a game of computer simulations and image matching, though it is a very huge and accurate game.

## References

[1] Abbott, B.P., *et al.* (2016) *Physical Review Letters*, **116**, 06112 1-16.

[2] Abbott, B.P., *et al.* (2016) *Physical Review Letters*, **116**, 241103 1-14.

[3] Liu, L. and Zhao, Z. (2004) General Relativity (Second Version). High Education Publishing Company, 140.

[4] Fang, H.L. (2014) Optical Resonant Cavity and Detection of Gravitational Waves. Science Publishing Company, 239, 246, 331.

[5] Mei, X. and Yu, P. (2017) *Journal of Modern Physics*, **7**, 1098-1104. http://dx.doi.org/10.4236/jmp.2016.710098

[6] Callen, H.B. and Green, R.F. (1952) *Physical Review Letters*, **86**, 702. http://dx.doi.org/10.1103/PhysRev.86.702

[7] Ohanian, H.C. and Ruffini, R. (1994) Gravitation and Space-Time. W. W. Norton & Company, Inc., 155.

[8] Cooperstock, C.F. and Faraoni, V. (1993) *Classical and Quantum Gravity*, **10**, 1989.

[9] Baroni, L., Fortini, P.L. and Gualdi, C. (1985) *Annals of Physics* (*New York*), **162**, 49.

[10] Moss, G.E., Miller, L.R. and Forward, R.L. (1971) *Applied Optics*, **10**, 2495-2498. http://dx.doi.org/10.1364/AO.10.002495

[11] Ulianov, P. (2016) *Global Journal of Physics*, **4**, 404-420.

Scientific Research Publishing

Journal of Modern Physics, 2016, 7, 1762-1776
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# The Inverse Gravity Inflationary Theory of Cosmology

**Edward A. Walker**

Mathematics Department, Florida Memorial University, Miami Gardens, FL, USA
Email: dwrdwalker@yahoo.com

**How to cite this paper:** Walker, E.A. (2016) The Inverse Gravity Inflationary Theory of Cosmology. *Journal of Modern Physics*, **7**, 1762-1776.
http://dx.doi.org/10.4236/jmp.2016.713158

**Received:** August 12, 2016
**Accepted:** September 25, 2016
**Published:** September 28, 2016



## Abstract

Cosmological expansion or inflation is mathematically described by the theoretical notion of inverse gravity whose variations are parameterized by a factor that is a function of the distance to which cosmological expansion takes prominence over gravity. This assertion is referred to as the inverse gravity inflationary assertion. Thus, a correction to Newtonian gravitational force is introduced where a parameterized inverse gravity force term is incorporated into the classical Newtonian gravitational force equation where the inverse force term is negligible for distances less than the distance to which cosmological expansion takes prominence over gravity. Conversely, at distances greater than the distance to which cosmological expansion takes prominence over gravity. The inverse gravity term is shown to be dominant generating universal inflation. Gravitational potential energy is thence defined by the integral of the difference (or subtraction) between the conventional Newtonian gravitational force term and the inverse gravity term with respect to radius ($r$) which allows the formulation, incorporation, and mathematical description to and of gravitational redshift, the Walker-Robertson scale factor, the Robinson-Walker metric, the Klein-Gordon lagrangian, and dark energy and its relationship to the energy of the big bang in terms of the Inverse gravity inflationary assertion. Moreover, the dynamic pressure of the expansion of a cosmological fluid in a homogeneous isotropic universe is mathematically described in terms of the inverse gravity inflationary assertion using the stress-energy tensor for a perfect fluid. Lastly, Einstein's field equations for the description of an isotropic and homogeneous universe are derived incorporating the mathematics of the inverse gravity inflationary assertion to fully show that the theoretical concept is potentially interwoven into the cosmological structure of the universe.

## Keywords

Isotropic and Homogeneous Universe, Inverse Gravity, Cosmological Inflation, Gravitational Redshift, Robertson-Walker Scale Factor, Klein-Gordon Lagrangian, Dark Energy, Stress-Energy Tensor, Friedman-Walker-Robertson Metric, Photon

## 1. Introduction

The theoretical notion that cosmological expansion or universal inflation occurs due to inverse variations in gravitational force whose rate of change is regulated by a limiting factor or parameter is introduced. Thus, it is asserted that cosmological expansion or inflation is an inherent property of nature mathematically described by the difference between conventional Newtonian gravitational force and its inverse term which is multiplied by an inflationary parameter which regulates its rate of change. The inflationary parameter multiplied by the inverse term of Newtonian gravitational force is determined by (and is a function of) an astronomical distance to which cosmological expansion over takes gravitational force on a cosmic scale. The establishment of the core concept of the inverse gravity inflationary assertion aforementioned is the foundation to describing the universe in terms of the new assertion. Thus, the aim of this paper is the incorporation of the inverse gravity inflationary assertion (IGIA) into proven and established mathematics describing cosmological inflation.

A more detailed introduction is that this paper formulates the correction to the Newtonian gravitational force equation incorporating an inverse gravity term and its conditions. This permits the derivation of gravitational potential energy in terms of the IGIA. Resultantly, the relationship between gravitational potential energy, dark energy, gravitational redshift, the Klein-Gordon Lagrangian, the energy of the big bang, the Robertson-Walker scale factor, and the Friedman-Walker-Robertson metric in terms of the inverse gravity inflationary assertion (IGIA) is formulated and defined. The IGIA is then applied to the stress-energy tensor for describing the dynamic pressure of an expanding cosmological fluid in a homogeneous and isotropic universe. Lastly, the IGIA is applied to Einstein's field equations for the description of a spherically homogeneous isotropic universe which establishes the inverse gravity inflationary assertion. This will elucidate how the IGIA is interwoven into the cosmological structure of the universe.

## 2. The Correction to the Newtonian Gravitational Force Equation and IGIA Inflationary Parameter

To begin the heuristic derivation, mass values $m_1$ and $m_2$ are the combined masses of cosmological bodies (such as galaxies) evenly dispersed over an isotropic and homogeneous universe and $G$ is the gravitational constant. The terms of gravitational force which are a function of radius $r$ are given such that [1]:

$$F_g(r) = \frac{Gm_1m_2}{r^2},\ F_g'(r) = B_0\left[\frac{r^2}{Gm_1m_2}\right] \tag{1.0}$$

where $F_g(r)$ is the classical expression of Newtonian gravitational force and $F_g'(r)$

is the inverse term of Newtonian gravitational force, constant $B_0$ is the inflationary factor or parameter. Inflationary factor $B_0$ is stated such that:

$$B_0 = \frac{1}{r_0^2}; 1 > B_0 > 0; r_0 > 1 \tag{1.01}$$

The constant $r_0$ is the astronomical distance to which cosmological expansion takes prominence over gravity. The inverse term of gravity $F'_g(r)$ can be re-expressed in terms of distance $r_0$ such that:

$$F'_g(r) = \left[\frac{1}{r_0^2}\right]\left[\frac{r^2}{Gm_1m_2}\right] \tag{1.02}$$

The total gravitational force $F_T(r)$ or the <u>Newtonian correction</u> is stated as the difference between force values $F'_g(r)$ and $F_g(r)$ such that:

$$F_T(r) = F'_g(r) - F_g(r) \tag{1.03}$$

Therefore,

$$F_T(r) = \left[\frac{1}{r_0^2}\right]\left[\frac{r^2}{Gm_1m_2}\right] - \frac{Gm_1m_2}{r^2} \tag{1.04}$$

The direction (+ or −) of the value of total force $F_T(r)$ has relationships defined by the inequalities of radius *r* and distance $r_0$ given by the conditions below.

$$\text{For } r > r_0; +F_T(r) \tag{1.05}$$

$$\text{For } r < r_0; -F_T(r) \tag{1.06}$$

Condition (1.05) describes cosmological expansion $+F_T(r)$ away from the gravitational force center or gravitational source for distances $r > r_0$.Conversely, for condition (1.06), inverse gravity term $F'_g(r)$ in total force $F_T(r)$ is negligible at distance $r < r_0$ causing force direction $-F_T(r)$ toward the center of gravitational force.

The value of the cosmological parameter of distance $r_0$ is determined where total force value $F_T(r)$ equals zero and radius *r* equals cosmological parameter $r_0$ ($r = r_0$ and $F_T(r_0) = 0$). Furthermore, this implies that for the condition of $F_T(r_0) = 0$, the force terms $F'_g(r_0)$ and $F_g(r_0)$ in total force $F_T(r_0)$ are equal ($F'_g(r_0) = F_g(r_0)$). Therefore, total force $F_T(r_0)$ can be presented such that:

$$F_T(r_0) = \left[\frac{1}{r_0^2}\right]\left[\frac{(r_0)^2}{Gm_1m_2}\right] - \frac{Gm_1m_2}{(r_0)^2} = 0 \tag{1.07}$$

This reduces to:

$$\left[\frac{1}{Gm_1m_2}\right] - \frac{Gm_1m_2}{(r_0)^2} = 0 \tag{1.08}$$

Solving Equation (1.08) for the cosmological parameter of distance $r_0$ gives a value such that:

$$r_0 = Gm_1m_2 \tag{1.09}$$

Conclusively Equation (1.09) above, gives the value of distance $r_0$ to which cosmological expansion takes prominence over gravity. In describing an isotropic and Homogeneous spherical universe, mass values $m_1$ and $m_2$ will be evenly (and uniformly) distributed about the spherical volume. Mass value $m_u$ denotes the total mass of the universe, therefore the dispersion of cosmological mass $m_u$ will be described via the function $g(m,\theta,\phi)$ at mass variable $m$ and the spherical coordinates at $\theta$ and $\phi$ [2]. Resultantly, mass $m'_u$ (corresponding to each mass value of $m_1$ and $m_2$) which represents a portion of cosmological mass dispersed about the sphere is set equal to function $g(m,\theta,\phi)$ as shown below.

$$m'_u = g(m,\theta,\phi) \tag{1.10}$$

Function $g(m,\theta,\phi)$ takes on a value of [2]:

$$g(m,\theta,\phi) = \left[(m\cos\theta\sin\phi)^2 + (m\sin\theta\sin\phi)^2 + (m\cos\phi)^2\right]^{1/2} \tag{1.11}$$

The gravitational interaction of symmetric portions of cosmological mass $m'_u$ and $m'_u$ separated by distance $r_0$ is stated as the product between the two mass values such that:

$$m'_u m'_u = \left(g(m,\theta,\phi)\right)\left(g(m,\theta,\phi)\right) = \left(g(m,\theta,\phi)\right)^2 \tag{1.12}$$

Thus, the continuous sums (or integration) of gravitational mass interaction $m'_u m'_u$ to whom are located on opposite sides of distance $r_0$ is taken to a value π and to the value of the mass of the universe $m_u$. Thus, the gravitational interaction of mass values $m_1 m_2$ in Equation (1.09) equals the triple integral shown below [2].

$$m_1 m_2 = \int_0^{m_u}\int_0^{\pi}\int_0^{\pi} m\left(m'_u m'_u\right)\mathrm{d}m\,\mathrm{d}\theta\,\mathrm{d}\phi \tag{1.13}$$

As the continuous sums of the integrals in Equation (1.13) progress in concert with angles $\theta$ and $\phi$ and sum up to a value $\pi$, the interacting mass values of $m_1$ and $m_2$ on opposite sides of distance $r_0$ sum up in concert with variable mass $m$ to encompass both halves of the spherical volume, giving the gravitational interaction of the entire spherical volume. Thus, the integration of (1.13) gives the gravitational mass interaction of the entire spherical volume. Therefore substituting the value of Equation (1.13) into (1.09) gives the proper mathematical description of distance $r_0$ shown below.

$$r_0 = G\left[\int_0^{m_u}\int_0^{\pi}\int_0^{\pi} m\left(m'_u m'_u\right)\mathrm{d}m\,\mathrm{d}\theta\,\mathrm{d}\phi\right] \tag{1.14}$$

The aim and scope of this paper is to introduce the notion and mathematics of the Inverse gravity inflationary assertion, thus we leave the calculation and value of Equation (1.14) as an exercise to the scientific community based on data obtained (The value of universal mass $m_u$) by astronomical observations.

## 3. The IGIA Mathematical Integration into Established Fundamental Concepts in Cosmology

This section applies the mathematical concept of the inverse gravity assertion to gravitational potential energy, gravitational Redshift, the Robertson-Walker scale factor, Friedman-Walker-Robertson metric, the Klein-Gordon lagrangian, dark energy, and the energy of the big bang. Gravitational potential energy $U_T(r)$ describing the energy of inflation in terms of the IGIA is equal to the conventional integral of total force $F_T(r)$ (Equation (1.04)) with respect to radius $r$ (for the condition of $r > r_0$) as shown below [1].

$$U_T(r) = \int F_T(r)\,\mathrm{d}r \tag{2.0}$$

Thus after evaluating the integral above, one obtains a value of potential energy $U_T(r)$ such that:

$$U_T(r) = \left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r} \tag{2.01}$$

As a photon propagates across the expanding cosmological expanse, its energy is affected by the gravitational potential energy $U_T(r)$. Thus, photonic energy $E$ is set equal to potential energy $U_T(r)$ in terms of the IGIA; this equivalence is displayed below [1].

$$E = pc = \frac{hc}{\lambda_g} = U_T(r) \tag{2.02}$$

where $\lambda_g$ is the photon's wavelength influenced by potential energy $U_T(r)$, the photonic energy affected by the potential energy field of $U_T(r)$ can be expressed such that [1]:

$$\frac{hc}{\lambda_g} = \left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r} \tag{2.03}$$

Resultantly, wavelength $\lambda_g$ affected by the potential energy field of $U_T(r)$ of the IGIA has a value expressed as [1]:

$$\lambda_g = hc\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right]^{-1} \tag{2.04}$$

Energy $E_0$ is the initial energy value ($E_0 = (hc/\lambda_0)$) of the emitted photon prior to it traversing through a region of space-time under the influence of gravitational potential energy $U_T(r)$ in terms of the IGIA, thus redshift $z$ is given such that [3]:

$$z = \frac{U_T(r) - E_0}{E_0} = \frac{\left(\dfrac{hc}{\lambda_g}\right) - E_0}{E_0} \equiv \frac{\lambda_0}{\lambda_g} - 1 \tag{2.05}$$

This reduces to [3]:

$$z = \frac{1}{E_0}\left(\frac{hc}{\lambda_g}\right) - 1 \equiv \frac{\lambda_0}{\lambda_g} - 1 \tag{2.06}$$

Red shift value $z$ can then be expressed in terms of the IGIA such that:

$$z = \frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right] - 1 \tag{2.07}$$

Thus, the photonic energy value $hc/\lambda_g$ is substituted by the value of potential energy $U_T(r)$ in Equation (2.06) giving the value of redshift $z$ in Equation (2.07). Observe the expression below where $a_0$ is the scale factor of the universe as it is presently and $a(t_{em})$ is the scale factor at the emission time $t_{em}$ of the photon (or a scale factor of the universe as it was in the past as some authors state it) [4].

$$1+z=\frac{a_0}{a(t_{em})} \tag{2.08}$$

Substituting the value of redshift $z$ of Equation (2.07) into (2.08) above gives [4]:

$$1+\left(\left(\frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right]+\frac{Gm_1m_2}{r}\right]\right)-1\right)=\frac{a_0}{a(t_{em})} \tag{2.09}$$

This reduces to:

$$\frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right]+\frac{Gm_1m_2}{r}\right]=\frac{a_0}{a(t_{em})} \tag{2.10}$$

The value of the scale factor at the time of the emitted photon $a(t_{em})$ is given such that [4]:

$$a(t_{em})=a_0\left[\frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right]+\frac{Gm_1m_2}{r}\right]\right]^{-1} \tag{2.11}$$

where Equation (2.11) is of the form $a(t)=1/(1+z)$ [4] which implies that $a_0$ equals 1 ($a_0=1$). Equation (2.11) adequately shows the relationship of the Robertson-Walker scale factors and gravitational redshift to the IGIA. At this juncture, the IGIA connection to Friedman-Lemaitre-Walker-Robertson metric $\mathrm{d}\Sigma^2$ is shown below [3].

$$\mathrm{d}\Sigma^2=-\mathrm{d}t^2+a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2}+r^2\mathrm{d}\theta^2+r^2\sin^2\theta\mathrm{d}\phi^2\right] \tag{2.12}$$

The scale factor $a(t)$ in Equation (2.12) above is set equal to the scale factor $a(t_{em})$ of Equation (2.11) ($a(t)=a(t_{em})$) giving Equation (2.12) such that (note: for our purposes $t=t_{em}$) [3]:

$$\mathrm{d}\Sigma^2=-\mathrm{d}t^2+a^2(t_{em})\left[\frac{\mathrm{d}r^2}{1-kr^2}+r^2\mathrm{d}\theta^2+r^2\sin^2\theta\mathrm{d}\phi^2\right] \tag{2.13}$$

Substituting the value of scale factor $a(t_{em})$ Equation (2.11) into Equation (2.13) above gives the Friedman-Walker-Robertson metric in terms of the IGIA such that:

$$\begin{aligned}\mathrm{d}\Sigma^2=&-\mathrm{d}t^2\\&+\left[a_0\left[\frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right]+\frac{Gm_1m_2}{r}\right]\right]^{-1}\right]^2\left[\frac{\mathrm{d}r^2}{1-kr^2}+r^2\mathrm{d}\theta^2+r^2\sin^2\theta\mathrm{d}\phi^2\right]\end{aligned} \tag{2.14}$$

This establishes the IGIA relationship to the Friedman-Walker-Robertson metric, where constant $k$ in Equation (2.14) is equal to +1 ($k=+1$) and $r>r_0$ for positive spherical curvature describing the expansion of the cosmological fluid [3].

The inverse gravity inflationary assertion (IGIA) can be defined in terms of field theory via its relationship to the Klein-Gordon lagrangian. In formulating the expressions describing this relationship, IGIA potential energy $U_T(r)$ is set equal relativistic energy denoted $E_{rel}$ as shown below [1].

$$U_T(r) = \left[p^2c^4 + m^2c^4\right]^{1/2} \equiv E_{rel} \tag{2.15}$$

This can be expressed such that:

$$E_{rel} = \left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right] = \left[p^2c^4 + m^2c^4\right]^{1/2} \tag{2.16}$$

where $\phi(x_{ц})$ is a scalar field function at Minkowski coordinates $x_{ц}$, the momenta $p$ is expressed as a tangent vector on the Minkowski coordinates $x_{ц}$ which is a function of (or parameterized by) time $t$ ( $x_{ц} = x_{ц}(t)$ ) as shown below [3].

$$p = \frac{1}{c}\frac{\partial\phi(x_{ц})}{\partial t} = \frac{1}{c}\frac{\partial\phi(x_{ц})}{\partial x_{ц}}\frac{\partial x_{ц}}{\partial t} = \frac{1}{c}\nabla_{ц}\phi(x_{ц}) \tag{2.17}$$

Expressing Equation (2.16) in terms of field $\phi(x_{ц})$ and substituting the value of momentum $p$ (of Equation (2.17)) into Equation (2.16) gives [5]:

$$E_{rel}(x_{ц}) \equiv \left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right] = \left[\left(\nabla_{ц}\phi(x_{ц})\right)^2 + \left(m^2\left(\phi(x_{ц})\right)^2c^4\right)\right]^{1/2} \tag{2.18}$$

where the relativistic energy $E_{rel}$ is expressed as a function of Minkowski coordinates $x_{ц}$ ( $E_{rel}(x_{ц})$ ) which gives a form of the Klein-Gordon equation. The speed of light $c$ is set equal to unity ( $c = 1$ ). A priori is that the differential momentum value $\nabla_{ц}\phi(x_{ц})$ relates to the energy value of the IGIA such that (or solving Equation (2.18) for the differential term $\nabla_{ц}\phi(x_{ц})$ ):

$$\nabla_{ц}\phi(x_{ц}) = \left[\left(E_{rel}(x_{ц})\right)^2 - \left(m^2\left(\phi(x_{ц})\right)^2\right)\right]^{1/2} \tag{2.19}$$

Equation (2.19) can be alternatively expressed in terms of the IGIA such that:

$$\nabla_{ц}\phi(x_{ц}) = \left[\left(\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right]\right)^2 - \left(m^2\left(\phi(x_{ц})\right)^2\right)\right]^{1/2} \tag{2.20}$$

where $r$ is a function of the Minkowski is coordinates ( $x_{ц} = \{it, x_1, x_2, x_3\}$ ), distance $r$ can be expressed such that [3]:

$$r = \left[-t^2 + \sum\nolimits_1^3 (x_{ц})^2\right]^{1/2} = r(x_{ц}) \tag{2.21}$$

Equation (2.19) can then be expressed in terms of radius $r(x_{ц})$ ( $r = r(x_{ц})$ ) as shown below.

$$\nabla_{ц}\phi(x_{ц}) = \left[\left(\left[\left[\frac{1}{r_0^2}\right]\left[\frac{\left(r(x_{ц})\right)^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r(x_{ц})}\right]\right)^2 - m^2\left(\phi(x_{ц})\right)^2\right]^{1/2} \tag{2.22}$$

Therefore, we introduce momentum $p'$ expressed as the differential term $\nabla^{ц}\phi\left(x_{ц}\right)$ (observe the superscript $ц$) as shown below [3].

$$p' = \nabla^{ц}\phi\left(x_{ц}\right) = \eta^{ц\sigma}\nabla_{ц}\phi\left(x_{ц}\right) = \eta^{ц\sigma}\frac{\partial\phi\left(x_{ц}\right)}{\partial t} \tag{2.23}$$

(Note: Recall thatthe speed of light $c$ is set equal to unity ($c = 1$)) Where the Minkowski metric $\eta^{ц\sigma}$ is expressed such that $\eta^{ц\sigma} = diag\left[-1,1,1,1\right]$ [3], relativistic energy $E'_{rel}\left(x_{ц}\right)$ corresponds to momentum $p'$ and the IGIA such that:

$$E'_{rel}\left(x_{ц}\right) \equiv \left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right] = \left[\left(\nabla^{ц}\phi\left(x_{ц}\right)\right)^2 + \left(m^2c^4\right)\right]^{1/2} \tag{2.24}$$

Equation (2.25) below is the Klein-Gordon equation expressed such that [3]:

$$\left(E_{KG}\phi\left(x_{ц}\right)\right)^2 = \nabla_{ц}\phi\left(x_{ц}\right)\nabla^{ц}\phi\left(x_{ц}\right) - m^2\left(\phi\left(x_{ц}\right)\right)^2 \tag{2.25}$$

Thus, as presented by Wald [3], the Klein-Gordon lagrangian is of the form shown below.

$$L_{KG} = \frac{1}{2}\left(E_{KG}\phi\left(x_{ц}\right)\right)^2 = -\frac{1}{2}\left[\nabla_{ц}\phi\left(x_{ц}\right)\nabla^{ц}\phi\left(x_{ц}\right) + m^2\left(\phi\left(x_{ц}\right)\right)^2\right] \tag{2.26}$$

Expressing the Klein-Gordon lagrangian (Equation (2.26)) above in terms of the IGIA, the value of Equation (2.23) is substituted into Equation (2.26) (where $\nabla^{ц}\phi\left(x_{ц}\right) = \eta^{ц\sigma}\nabla_{ц}\phi\left(x_{ц}\right)$) giving:

$$L_{KG} = -\frac{1}{2}\left[\eta^{ц\sigma}\left[\left(\left[\left[\frac{1}{r_0^2}\right]\left[\frac{\left(r\left(x_{ц}\right)\right)^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r\left(x_{ц}\right)}\right]\right)^2 - m^2\left(\phi\left(x_{ц}\right)\right)^2\right] + m^2\left(\phi\left(x_{ц}\right)\right)^2\right] \tag{2.27}$$

This implies that the product of differential terms $\nabla_{ц}\phi\left(x_{ц}\right)\nabla^{ц}\phi\left(x_{ц}\right)$ takes a value incorporating the IGIA such that:

$$\nabla_{ц}\phi\left(x_{ц}\right)\nabla^{ц}\phi\left(x_{ц}\right) = \eta^{ц\sigma}\left[\left(\left[\left[\frac{1}{r_0^2}\right]\left[\frac{\left(r\left(x_{ц}\right)\right)^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r\left(x_{ц}\right)}\right]\right)^2 - m^2\left(\phi\left(x_{ц}\right)\right)^2\right] \tag{2.28}$$

Solutions to Equations (2.27) and (2.28) pertain to mathematical methods of solving differential equations. Conclusively, Equation (2.28) is the IGIA correlation to various areas of field theory especially quantum energy fields describing vacuum energy (and the stress-energy tensor described in the next section). Lastly, dark energy $E_{dark}$ in terms of the IGIA is given by the conditions shown below.

$$E_{dark} = \left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] \tag{2.29}$$

Thus, dark energy is interpreted according to the IGIA as the inverse term of potential energy $U_T\left(r\right)$ (of Equation (2.01)). An important consideration is that the energy of expansion or energy $E_{dark}$ is different from the energy of the big bang which will be denoted energy $E_{BB}$ for the purposes of this explanation. Thus energy $E_{BB}$ is far greater in magnitude than the energy of gravity at the big bang denoted $U_{BB} = \left( GMm/r_i \right)$ (along with other elementary forces such as electromagnetism, the strong nuclear force, and the weak nuclear force opposing universal expansion) at the big bang where radius $r_i$ is the infinitesimally small distance of the big bang ($E_{BB} > U_{BB}$). Where big bang energy $E_{BB}$ is composed of kinetic energy and electromagnetic energy of the universe, this implies that energy $E_{BB}$ is also sufficient for accelerating the total cosmological mass beyond the astronomical distance or radius $r_0$ to which cosmological expansion takes prominence over gravity and where energy $E_{dark}$ can generate cosmological expansion.

## 4. The Dynamic Pressure of an Expanding Cosmological Fluid in Terms of the IGIT

This section mathematically defines the dynamic pressure of a cosmological fluid in a homogeneous isotropic universe in terms of the IGIA. The stress-energy tensor for a perfect fluid used to describe the expansion of the cosmological fluid is given such that [3]:

$$T_{ab} = \nabla_a \boldsymbol{\phi} \nabla_b \boldsymbol{\phi} + g_{ab} L_{KG} \tag{3.0}$$

where $\boldsymbol{\phi}$ is the field function of the space-time manifold [5] that the stress-energy tensor is defined on, $g_{ab}$ is the metric tensor, and $L_{KG}$ is the Klein-Gordon lagrangian such that [3]:

$$L_{KG} = -\frac{1}{2}\left[ \nabla^c \boldsymbol{\phi} \nabla_c \boldsymbol{\phi} + \boldsymbol{\phi}^2 m^2 c^2 \right] \tag{3.01}$$

The 4 space tangent vectors $\nabla_a \boldsymbol{\phi} \nabla_b \boldsymbol{\phi}$ in Equation (3.0) obey the geodesic rule such that [3]:

$$\nabla_c \left( \nabla_a \boldsymbol{\phi} \nabla_b \boldsymbol{\phi} \right) = \nabla_c \left( \nabla_a \boldsymbol{\phi} \nabla_b \boldsymbol{\phi} \right) + \boldsymbol{\Gamma}^c_{ab} \left( \nabla_a \boldsymbol{\phi} \nabla_b \boldsymbol{\phi} \right) \equiv 0 \tag{3.02}$$

Thus showing the appropriate use of the Christoffle symbol $\boldsymbol{\Gamma}^c_{ab}$ [3]. The partial derivative $\nabla_c$ in Equation (3.02) is of the form $\nabla_c = \left( \partial / \partial x_c \right)$. The tangent vector of the form $\nabla_\mu \boldsymbol{\phi}$ is defined by the chain rule such that [2]:

$$\nabla_\mu \boldsymbol{\phi} = \frac{\partial \boldsymbol{\phi}}{\partial x_{\text{ц}}} \frac{\partial x_{\text{ц}}}{\partial t} \tag{3.03}$$

where the time coordinate has a value $ct$ and the speed of light $c$ is set to unity ($c = 1$) [5], the spatial coordinates $x_{\text{ц}}$ in $R^4$ are the Minkowski coordinates ($x_{\text{ц}} = \left\{ it, x_1, x_2, x_3 \right\}$ and $x_{\text{ц}} = x_{\text{ц}}(t)$) [5]. Thus the tangent vector $\nabla_\mu \boldsymbol{\phi}$ at time $t$ is the <u>4-velocity of the cosmological fluid</u> denoted $u_{\text{ц}}$ ($u_{\text{ц}} \in R^4$) as shown below [3].

$$u_{\text{ц}} = \nabla_\mu \boldsymbol{\phi} = \frac{\partial \boldsymbol{\phi}}{\partial x_{\text{ц}}} \frac{\partial x_{\text{ц}}}{\partial t} \tag{3.04}$$

Therefore the product of tangent vectors $\nabla_a \boldsymbol{\phi}$ and $\nabla_b \boldsymbol{\phi}$ ($\nabla_a \boldsymbol{\phi} \nabla_b \boldsymbol{\phi}$) are given such that [3]:

$$\nabla_a \boldsymbol{\phi} \nabla_b \boldsymbol{\phi} = u_a \cdot u_b = \sum\nolimits_{\text{ц}} \left( u_{\text{ц}} \right)^2 \tag{3.05}$$

It must be noted that tangent vectors $\nabla_a\phi$ and $\nabla_b\phi$ are symmetric ($\nabla_a\phi = \nabla_b\phi$) which implies the fluid velocities $u_a$ and $u_b$ are also symmetric ($u_a = u_b = u_\mu$), therefore $u_a \cdot u_b = \sum_\mu (u_\mu)^2$. Total dynamic pressure $P_{ab}$ (where the subscripts *ab* pertain to a 4 by 4 matrix) is given in terms of fluid 4-velocity $u_\mu$ such that [1] [3]:

$$P_{ab} = \frac{\rho \sum_\mu (u_\mu)^2}{2} \tag{3.06}$$

This implies that [1]:

$$P_{ab} = \frac{\rho \sum_\mu (u_\mu)^2}{2} = \frac{\rho(\nabla_a\phi\nabla_b\phi)}{2} \tag{3.07}$$

In the task of defining the expansion of the cosmological fluid in terms of the IGIT, consider the unit vector $\hat{u}$ in $R^4$ ($\hat{u} \in R^4$ and $x_\mu \in R^4$) shown below [2].

$$\hat{u} = \frac{x_\mu}{|x_\mu|} \quad \text{such that } \hat{u} \cdot \hat{u} = 1 \tag{3.08}$$

Multiplying unit vector $\hat{u}$ to the IGIT force value $F_T(r)$ of Equation (1.03) gives vector valued force $(F_T(r))_\mu$ in $R^4$ ($(F_T(r))_\mu \in R^4$) such that [2]:

$$(F_T(r))_\mu = \hat{u}(F_T(r)) \tag{3.09}$$

This can be expressed such that:

$$(F_T(r))_\mu = \frac{x_\mu}{|x_\mu|}\left[F_g'(r) - F_g(r)\right] \tag{3.10}$$

where $A'$ ($A' = 4\pi r^2$ [3]) is a spherically symmetric area, the sum or superposition of 4 pressure components is as expressed below [3]. [Where pressure = force ÷ area ($P = F/A$)][1]

$$\sum_\mu \frac{(F_T(r))_\mu}{A'} = \sum_\mu \left[\frac{x_\mu}{A'|x_\mu|}\left[F_g'(r) - F_g(r)\right]\right] \to \sum \frac{F}{A} \tag{3.11}$$

(Note: that the pressure component $P_{00}$ of the 4 by 4 matrix of the stress-energy tensor has a value of energy density $\rho_E$ ($P_{00} = \rho_E$) [3]) This is set equal to total dynamic pressure $P_{ab}$ of Equation (3.07) as shown below [1].

$$P_{ab} = \frac{\rho \sum_\mu (u_\mu)^2}{2} = \sum_\mu \frac{(F_T(r))_\mu}{A'} \tag{3.12}$$

This implies that [1]:

$$\frac{\rho \sum_\mu (u_\mu)^2}{2} = \sum_\mu \frac{(F_T(r))_\mu}{A'} = \frac{\rho(\nabla_a\phi\nabla_b\phi)}{2} \equiv P_{ab} \tag{3.13}$$

which also implies the equivalence of:

$$\nabla_a\phi\nabla_b\phi = \frac{2}{\rho}\sum_\mu \frac{(F_T(r))_\mu}{A'} \equiv \frac{2}{\rho}P_{ab} \tag{3.14}$$

Therefore this can be expressed such that:

$$\nabla_a\phi\nabla_b\phi = \frac{2}{\rho}\sum_\mu\left[\frac{x_\mu}{A'|x_\mu|}\left[F'_g(r)-F_g(r)\right]\right] \equiv \frac{2}{\rho}P_{ab} \tag{3.15}$$

Now substituting the value of force term $F'_g(r)-F_g(r)$ into Equation (3.15) gives:

$$\nabla_a\phi\nabla_b\phi = \frac{2}{\rho}\sum_\mu\left[\frac{x_\mu}{A'|x_\mu|}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^2}{Gm_1m_2}\right]-\frac{Gm_1m_2}{r^2}\right]\right] \equiv \frac{2}{\rho}P_{ab} \tag{3.16}$$

Thence, substituting the value of $\nabla_a\phi\nabla_b\phi$ (shown above) for the value of Equation (3.15) into the stress-energy tensor $T_{ab}$, the stress-energy tensor can be expressed such that:

$$T_{ab} = \frac{2}{\rho}\sum_\mu\left[\frac{x_\mu}{A'|x_\mu|}\left[F'_g(r)-F_g(r)\right]\right] - g_{ab}L_{KG} \tag{3.17}$$

which is equivalent to the stress-energy tensor such that:

$$T_{ab} = \frac{2}{\rho}\sum_\mu\left[\frac{x_\mu}{A'|x_\mu|}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^2}{Gm_1m_2}\right]-\frac{Gm_1m_2}{r^2}\right]\right] - g_{ab}L_{KG} \tag{3.18}$$

Equation (3.18) can be expressed in matrix form such that:

$$T_{ab} = \begin{bmatrix} \rho_E & T_{01} & T_{02} & T_{03} \\ T_{10} & \left(\frac{2}{\rho}P_{11}-g_{11}L_{KG}\right) & T_{12} & T_{13} \\ T_{20} & T_{21} & \left(\frac{2}{\rho}P_{22}-g_{22}L_{KG}\right) & T_{23} \\ T_{30} & T_{31} & T_{32} & \left(\frac{2}{\rho}P_{33}-g_{33}L_{KG}\right) \end{bmatrix}$$

The correlation of the stress-energy tensor $T_{ab}$ formulated in terms of the IGIA to Einstein's field equations describing an isotropic and homogeneous universe is given in the conclusion. Thus, we conclude with the formulation and incorporation of the IGIA in Einstein's field equation in its entirety.

## 5. Conclusions: Einstein's Field Equations Describing an Expanding Homogeneous and Isotropic Universe in Terms of the IGIA

In describing an expanding homogeneous isotropic universe in terms of the IGIA, it is of great importance that the IGIA is fully incorporated to Einstein's field equations as a whole. Thus, we began the heuristic derivation according to Wald [3] with expressions of Einstein's Field equations such that:

$$G_{tt} = 8\pi T_{tt} = 8\pi\rho \tag{4.0}$$

$$G_{**} = 8\pi T_{**} = 8\pi P \tag{4.1}$$

The expressions of the stress-energy tensor in Equations (4.0) and (4.10 ($T_{tt}$ and $T_{**}$) are related the average value of cosmological mass density $\rho$ and to pressure value $P$

[3]. The two expressions of Einstein's tensor $G_{tt}$ and $G_{**}$ are given such that [3]:

$$G_{tt} = G_{ab}u^a u^b \tag{4.2}$$

$$G_{**} = G_{ab}s^a s^b \tag{4.3}$$

where $u^{\text{ц}}$ and $s^{\text{ц}}$ are contra variant unit tangent vectors of homogenous surfaces of the isotropic and expanding universe [3]. Thus as depicted by Wald, the Robertson-Walker metric for describing a flat spatial geometry is expressed such that [3]:

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right) \tag{4.4}$$

The scale factor at time *t* denoted $a(t)$ in the metric above is equal to the value of scale factor $a(t_{em})$ of Equation (2.11) ( $a(t) = a(t_{em})$ ) in terms of the IGIA. This equivalence can be stated such that:

$$a(t) = a(t_{em}) = a_0\left[\frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right]\right]^{-1} \tag{4.5}$$

Recall that for our purposes $t = t_{em}$. Substituting the value of Equation (4.5) into Equation (4.4), the flat space Robertson-Walker metric of equation of (4.4) can be stated in terms of the IGIA (similarly to Robertson-Walker metric of Equation (2.14)) such that:

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \left[a_0\left[\frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right]\right]^{-1}\right]^2\left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right) \tag{4.6}$$

Thus solving Equation (4.6) for the value of scale factor $a(t_{em})$ in Equation (4.6) gives the equivalence of values:

$$a(t_{em}) = \left[\left(\mathrm{d}s^2 + \mathrm{d}t^2\right)\left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right)^{-1}\right]^{\frac{1}{2}} = a_0\left[\frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^3}{3Gm_1m_2}\right] + \frac{Gm_1m_2}{r}\right]\right]^{-1} \tag{4.7}$$

The time derivative of scale factor $a(t_{em})$ is denoted $\dot{a}$ (and *a* is simply $a = a(t_{em})$) and can be expressed such that [3]:

$$\dot{a} = \frac{\partial a(t_{em})}{\partial t} \tag{4.8}$$

The coordinate $x_{\text{ц}}$ represent the Minkwoski coordinates ( $x_{\text{ц}} = \{it, x, y, z\}$ ) [5]. Where the time coordinate has a value of *ct* and the speed of light *c* is set to unity ( $c = 1$ ) [5]. The question in reference to calculations is "how does the scale factor $a(t_{em})$ in terms of the IGIA (Equation (2.11) or (4.5)) relate to the time valued-derivative $\dot{a}$ of Equation (4.8) shown above?". The radius *r* is defined on the Minkowski coordinates ( $it, x, y, z$ ), thus the value (or magnitude) of radius *r* (where $r^2 = \Delta s^2$ ) is expressed such that [5]:

$$\begin{aligned} r^2 = \Delta s^2 &= \left[(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2\right]^{1/2} \\ &= \left[(it - it_i)^2 + (x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2\right]^{1/2} \end{aligned} \tag{4.9}$$

Distance $r$ is measured from the center of expansion (or the center of the universe), thus the initial values of $t_i$, $x_i$, $y_i$, and $z_i$ equal zero. Therefore substituting zero for all initial values $t_i$, $x_i$, $y_i$, and $z_i$ ($(t_i, x_i, y_i, z_i) = (0,0,0,0)$) and solving for radius $r$ in Equation (4.9) gives (Similarly to Equation (2.20)):

$$r = \left[-t^2 + x^2 + y^2 + z^2\right]^{1/2} \tag{4.10}$$

Substituting the value of Equation (4.10) into Equation (2.11) (or Equation (4.5)) gives the IGIA scale factor as a function of the Minkowski coordinates ($a(t,x,y,z)$) such that:

$$\begin{aligned} a(t_{em}) &= a(t,x,y,z) \\ &= a_0\left[\frac{1}{E_0}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{\left(-t^2+x^2+y^2+z^2\right)^{3/2}}{3Gm_1m_2}\right]+\frac{Gm_1m_2}{\left[-t^2+x^2+y^2+z^2\right]^{1/2}}\right]\right]^{-1} \end{aligned} \tag{4.11}$$

Thus pertaining to the time coordinate ($x_0 = it$), the scale factor $a(t_{em})$ in terms of the IGIA is now differentiable to the time valued derivative of Equation (4.8), therefore permitting the continuation of the formulation without ambiguity. Equation (4.11) affords the opportunity to briefly present Hubble's constant in terms of the IGIA such that [3]:

$$H(t) = \frac{\dot{a}}{a} = \left[\frac{1}{a(t,x,y,z)}\right]\frac{\mathrm{d}a(t,x,y,z)}{\mathrm{d}t} \tag{4.12}$$

In continuing the IGIA's incorporation to Einstein's field equation, the scale factors $a$ (keep in mind that $a = a(t) = a(t_{em}) = a(t,x,y,z)$) relate to the symmetric Christoffel symbols such that [3]:

$$\Gamma^t_{xx} = \Gamma^t_{yy} = \Gamma^t_{zz} = a\dot{a} \tag{4.13}$$

$$\Gamma^x_{xt} = \Gamma^x_{tx} = \Gamma^y_{ty} = \Gamma^y_{yt} = \Gamma^z_{zt} = \Gamma^z_{tz} = \dot{a}/a \tag{4.14}$$

Thus we acknowledge that the Christoffel symbols $\Gamma^a_{bc}$ is of the form [3]:

$$\Gamma^a_{bc} = \frac{1}{2}\sum\nolimits_d g^{ad}\left\{\frac{\partial g_{cb}}{\partial x^b} + \frac{\partial g_{ca}}{\partial x^c} - \frac{\partial g_{bc}}{\partial x^c}\right\} \tag{4.15}$$

The components of the Ricci tensor are calculated according to the equation of [3]:

$$R_{ab} = \sum\nolimits_c R_{acb^c} \tag{4.16}$$

This can alternatively be expressed such that [3]:

$$R_{ab} = \sum\nolimits_c \frac{\partial y}{\partial x^c}\Gamma^c_{ab} - \frac{\partial}{\partial x^a}\left(\sum\nolimits_c \Gamma^c_{cb}\right) + \sum\nolimits_{d,c}\left(\Gamma^d_{ab}\Gamma^c_{dc} - \Gamma^d_{cb}\Gamma^c_{da}\right) \tag{4.17}$$

The Ricci tensor is then related to the scale factor $a$ (or $a(t_{em})$ in terms of the IGIA) by the equations of (where $\ddot{a} = \mathrm{d}^2\left(a(t,x,y,z)\right)/\mathrm{d}t^2$) [3]:

$$R_{tt} = -3\ddot{a}/a \tag{4.18}$$

$$R_{**} = a^{-2}R_{xx} = \frac{\ddot{a}}{a} + 2\frac{(\dot{a})^2}{a^2} \tag{4.19}$$

As stated by Wald, the value of Ricci tensor $R_{xx}$ in Equation (4.19) above relates to the Christoffel symbol such that [3]:

$$R_{xx} = \sum\nolimits_c \frac{\partial y}{\partial x^c} \Gamma^c_{xx} - \frac{\partial}{\partial x^x}\left(\sum\nolimits_c \Gamma^c_{cx}\right) + \sum\nolimits_{d,c}\left(\Gamma^d_{xx}\Gamma^c_{dc} - \Gamma^d_{cx}\Gamma^c_{dx}\right) \tag{4.20}$$

Therefore the value of the scalar curvature $R$ is given such that [3]:

$$R = -R_{tt} + 3R_{**} \tag{4.21}$$

Substituting the value of Equation (4.18) and (4.19) into Equation (4.21) give a value such that [3]:

$$R = -R_{tt} + 3R_{**} = 6\left(\frac{\ddot{a}}{a} + \frac{(\dot{a})^2}{a^2}\right) \tag{4.22}$$

Conclusively, the values of Einstein tensor values $G_{tt}$ and $G_{**}$ are given such that [3]:

$$G_{tt} = R_{tt} + \frac{1}{2}R = \frac{3(\dot{a})^2}{a^2} = 8\pi\rho \tag{4.23}$$

$$G_{**} = R_{**} - \frac{1}{2}R = -2\frac{\ddot{a}}{a} - \frac{(\dot{a})^2}{a^2} = 8\pi P \tag{4.24}$$

As stated by Wald, using Equation (4.23), Equation (4.24) can be rewritten such that [3]:

$$\frac{3(\dot{a})^2}{a^2} = -4\pi(\rho + 3P) \tag{4.25}$$

Due to the fact that the description of the IGIA is defined in reference to a homogeneous and isotropic universe, the general evolutions for isotropic and homogeneous universe are given such that [3]:

$$\frac{3(\dot{a})^2}{a^2} = 8\pi\rho - \frac{3k}{a^2} \tag{4.26}$$

$$\frac{3\ddot{a}}{a} = --4\pi(\rho + 3P) \tag{4.27}$$

where scale factors $a$ and their corresponding time derivatives ($\dot{a}$ and $\ddot{a}$) can be defined in terms of the IGIA of Equation (4.10) ($a = a(t, x, y, z)$), constant $k$ is equal to +1 ($k = +1$) and $r > r_0$ for positive spherical curvature describing the expansion of the cosmological fluid in an homogeneous isotropic universe [3]. At this juncture, the relationship of the scale factor of Equation (4.5) (and Equation (2.11)) in terms of the IGIA to Einstein's field equation have been formulated.

Equation (4.1) ($G_{**} = 8\pi T_{**} = 8\pi P$) shows the relationship between pressure $P$ and the stress-energy tensor $T_{ab}$ of Equation 3.18 [3]. This implies that $P$ can be minimally substituted for the stress-energy tensor ($P \to T_{ab}$). Thus, the pressure term $P$ relates the IGIA stress-energy tensor of Equation (3.18) (of the previous section) such that:

$$P \to T_{ab} = \frac{2}{\rho}\sum\nolimits_\mu \left[\frac{x_\mu}{A'|x_\mu|}\left[\left[\frac{1}{r_0^2}\right]\left[\frac{r^2}{Gm_1m_2}\right] - \frac{Gm_1m_2}{r^2}\right]\right] - g_{ab}L_{KG} \tag{4.28}$$

Resultantly, this can be expressed such that:

$$P = \frac{2}{\rho} \sum_{\mu} \left[ \frac{x_\mu}{A' |x_\mu|} \left[ \left[ \frac{1}{r_0^2} \right] \left[ \frac{r^2}{Gm_1 m_2} \right] - \frac{Gm_1 m_2}{r^2} \right] \right] - g_{ab} L_{KG} \tag{4.29}$$

Spherically symmetric area $A'$ is equal to a value of 8π ( $A' = 8\pi$ ), therefore Equation (4.29) can be stated such that:

$$P = \frac{2}{8\pi\rho} \sum_{\mu} \left[ \frac{x_\mu}{|x_\mu|} \left[ \left[ \frac{1}{r_0^2} \right] \left[ \frac{r^2}{Gm_1 m_2} \right] - \frac{Gm_1 m_2}{r^2} \right] \right] - g_{ab} L_{KG} \tag{4.30}$$

Substituting the value of pressure $P$ presented above into Equation (4.27) gives a value such that:

$$\frac{3\ddot{a}}{a} = -4\pi \left( \rho + 3 \left[ \frac{2}{8\pi\rho} \sum_{\mu} \left[ \frac{x_\mu}{|x_\mu|} \left[ \left[ \frac{1}{r_0^2} \right] \left[ \frac{r^2}{Gm_1 m_2} \right] - \frac{Gm_1 m_2}{r^2} \right] \right] - g_{ab} L_{KG} \right] \right) \tag{4.31}$$

Equation (4.31) gives an additional incorporation of the Mathematics of the IGIA showing that the theoretical concept is well ingrained to the cosmological structure of the universe. The incorporation of the IGIA mathematics to Einstein's field equations gives a complete description to validate the concept and convey a new theoretical possibility to the physics community.

## Acknowledgements

I would like to thank Juanita Walker and Mr. Eugene Thompson for their encouragement and support to this endeavor.

## References

[1] Young, H.D., Freedman, R.A. (2004) Sears and Zemansky's University Physics. 11th Edition, Addison Wesley, San Francisco, Cal., Pearson, 444, 445, 454-457, 517, 529, 530, 1431, 1450.

[2] Stewart, J. (2003) Calculus. 5th Edition, Belmont, Cal., Thompson, Brooks/Cole, 877, 968, 1095.

[3] Wald, R.M. (1984) General Relativity. Chicago Press, Ltd., Chicago, 17, 23-26, 31-38, 59, 70, 72, 95, 96, 101-104, 120-124.

[4] Mukhanov, V.F. (2005) Physical Foundations of Cosmology. Cambridge University Press, Cambridge, New York, Page 58. http://dx.doi.org/10.1017/CBO9780511790553

[5] Penrose, R. (2004) The Road to Reality: Comprehensive Guide to the Laws of the Universe. Alfred A. Knopf Publishing, New York, 222-223, 412, 413.

Journal of Modern Physics, 2016, 7, 1777-1800
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Cosmology of the Nambu-Jona-Lasinio Model

**Leonardo Quintanar G., A. de la Macorra**

Instituto de Fisica, Universidad Nacional Autonoma de Mexico, A.P. 20-364, 01000, Mexico D.F., Mexico
Email: a.macorra@gmail.com

**How to cite this paper:** Leonardo Quintanar G. and de la Macorra, A (2016) Cosmology of the Nambu-Jona-Lasinio Model. *Journal of Modern Physics*, **7**, 1777-1800.
http://dx.doi.org/10.4236/jmp.2016.713159

**Received:** May 3, 2016
**Accepted:** September 27, 2016
**Published:** September 30, 2016



## Abstract

We review the Nambu and Jona-Lasinio model (NJL), proposed long time ago, in the sixties, as a fermion interaction theory with chiral symmetry. The theory is not renormalizable and presents a symmetry breaking due to quantum effects which depends on the strength of the coupling constant. We may associate a phase transition with this symmetry breaking, leading from fermion states to a fermion condensate which can be described effectively by a scalar field. Our purpose in this paper is to exploit the interesting properties of NJL in a different context other than particle physics by studying its cosmological dynamics. We are interested in finding whether possibly the NJL model could be used to describe the still unknown dark energy and/or dark matter, from up to 95% of the energy content of the universe at present time.

## Keywords

Cosmology, Nambu Jona Lasinio

## 1. Introduction

In the last years the study of our universe has received a great deal of attention since, on the one hand fundamental theoretical cosmological questions remain unanswered and, on the other hand we have now the opportunity to measure the cosmological parameters with an extraordinary precision. In the last decades, research in cosmology has revealed the presence of unexplained forms of matter and energy called Dark Energy "DE" and Dark Matter "DM" making up to 95% of the energy content of the universe at present time. The study of supernovas SNIa shows that the universe is not only expanding, but besides it is accelerating [1]-[6]. Such behaviour can be explained by the existence of a new form of energy, Dark Energy with an anti-gravitational property, which would be explained by a fluid with negative pressure. Independent evidence for Dark Matter (DM) and Dark Energy (DE), is provided through the analysis of the

Cosmic Microwave Background radiation (CMB) [7]-[10], which has been measured by satellite WMAP [11], and more recently by Planck mission [12], and the dynamics of galaxies, clusters and super clusters, and the study of the formation of Large Scale Structure [13]-[16] in the universe and weak lensing (the gravitational deviation of light), which point out the existence of matter that do interacts with ordinary standard model matter only weakly, as due to gravity. Other important measurements are the Baryon Acoustic Oscillations "BAO" [17]-[19].

It has been established that our universe is flat and dominated at present time by Dark Energy "DE" and Dark Matter "DM" with $\Omega_{DE} \simeq 0.692 \pm 0.02$, $\Omega_m = 0.308 \pm 0.009$ and Hubble constant $H_o = (67.27 \pm 0.66)$ $\mathrm{km \cdot s^{-1} \cdot Mpc^{-1}}$ [12]. However, the nature and dynamics of Dark Energy and Dark Matter are topics of major interest in the field [20]. The equation of state "EOS" of DE is at present time $w_o \simeq -0.93 \pm 0.13$ but we still do not have a precise measurement of $w(z)$ as a function of redshift $z$ [12] [16]. Since the properties of Dark Energy are still under investigation, different DE parametrizations have been proposed to help discern on the dynamics of DE [20]-[23]. Some of these DE parametrizations have the advantage of having a reduced number of parameters, but they may lack a physical motivation and may also be too restrictive. Perhaps the best physically motivated candidates for Dark Energy are scalar fields which can be minimally coupled, only via gravity, to other fluids [20]-[23] or can interact weakly in interacting Dark Energy "IDE" [24]-[27]. Scalar fields have been widely studied in the literature [20]-[23] and special interest was devoted to tracker fields [22] [23] since in this case the behavior of the scalar field $\phi$ is very weakly dependent on the initial conditions at a very early epoch and well before matter-radiation equality. In this class of models the fundamental question of why DE is relevant now, also called the coincidence problem, can be ameliorated by the insensitivity of the late time dynamics on the initial conditions of $\phi$.

Nowadays there are a huge number of ideas aimed to explain these unknown cosmological fluids DE and DM, from the theoretical point of view, none of them being still conclusive. This situation supports and motivates our research. Given that our most successful theory of matter, the Standard Model of particle physics (SM), which is settled within the theoretical frame of Quantum Field Theory (QFT), it would be reasonable to ask a theory attempting to describe dark fluids to be based on QFT as well. In this paper we study a fermion interaction theory with a chiral symmetry, the Nambu-Jona-Lasinio (NJL) model. Though this is an old and well known model in the context of hadron physics, it has interesting properties and it is worth to consider it with a different perspective, by studying its possible relevance for Cosmological Physics. Other examples of QFT models of DE and DM have been proposed using gauge groups, similar to QCD in particle physics, and have been studied to understand the nature of Dark Energy [28] [29] and also Dark Matter [30] [31].

We organized the present work as follows: In Section 2 we present the NJL model. In Section 3 we review the pertinent cosmological theory. Sections 4 and 5 present a study of the cosmological dynamics of a NJL fluid with a weak and strong coupling, respectively. In Section 6 we consider the addition of a cosmological constant to our NJL fluid, and analyze the different possible behaviours. In Section 7 we comment an interesting possible way to modify the original NJL model, obtaining an additional term in the effective potential which could be related with a Cosmological Constant. Finally, in Section 8 we summarize our results and present the conclusions.

## 2. The Nambu-Jona-Lasinio Model

Inspired by a, by then recently explained phenomenon in Superconductivity research, professors Y. Nambu and Jona-Lasinio, suggested that the mass of fermion particles (described by a Dirac equation) could be generated from a primary four-fermion self interaction, leading to a chiral symmetry breaking. The proposed Lagrangian, invariant under chiral transformations, has the form

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi + \frac{g^2}{2}\left[\left(\bar{\psi}\psi\right)^2 - \left(\bar{\psi}\gamma_5\psi\right)^2\right], \tag{1}$$

where $\psi$ is a four-component spinor, and $g$ is a coupling constant. From Equation (1) the four fermion interaction term is given by

$$\mathcal{L}_{int} = \frac{g^2}{2}\left[\left(\bar{\psi}\psi\right)^2 - \left(\bar{\psi}\gamma_5\psi\right)^2\right] \tag{2}$$

with no original mass term for the fermions. Since the coupling has dimension-2 in mass units, the theory is non-renormalizable. However, we are interested in considering the NJL model as an *effective theory*, useful below certain energy scale. The theory (1) describes a four-fermion interaction which can be expanded following conventional perturbation theory, and represented by Feynman diagrams (**Figure 1**).

The infinite number of fermion loops can be resumed giving a non-perturbative potential. This can be easily done by introducing an auxiliary scalar field $\phi$ and an equivalent Lagrangian for Equation (2) in the form

$$\mathcal{L}_{int} = mg\varphi\bar{\psi}\psi - \frac{1}{2}m^2\phi^2. \tag{3}$$

The field $\phi$ plays the role of a Lagrange multiplier which can be eliminated using the Euler-Lagrange equations, $\partial_{\mu}\frac{\delta\mathcal{L}}{\delta\left(\partial_{\mu}\phi\right)} - \frac{\delta\mathcal{L}}{\delta\phi} = 0$. For the Lagrangian above we find

$$\phi = \frac{g}{m}\bar{\psi}\psi, \tag{4}$$

where $\phi$ has mass dimensions, and by substituting Equation (4) in Equation (3) one can recover the original Lagrangian Equation (2). Note that we introduced the parame-

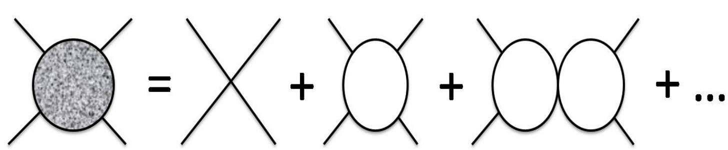



**Figure 1.** Feynman diagram for a four-fermion interaction.

ter $m$ with a physical dimension of mass, so that $mg$ is a dimensionless coefficient, and we have dimensional consistency for all the physical quantities.[1] The term $\bar{\psi}\gamma_5\psi$ in Equation (1) represents a pseudo-scalar quantity, and we have allowed ourselves to ignore the field contribution associated with it in the new Lagrangian in Equation (3), as we would like to start to study the simplest possible model.[2]

From the equivalent Lagrangian one may read the fermion mass and the tree level scalar potential $V_0$. We have respectively:

$$m_\psi^2 = \left(mg\phi\right)^2, \quad V_0 = \frac{1}{2}m^2\phi^2. \tag{5}$$

The effect of quantum processes (represented by loop diagrams) may be taken into account through the well known Coleman-Weinberg potential

$$V_1 = -\frac{1}{8\pi}\int p^2 \log\left(p^2 + m_\psi^2\right)\mathrm{d}^2 p, \tag{6}$$

the minus sign in $V_1$ is because it corresponds to the fermionic contribution to the Coleman-Weinberg potential. As we will see, for the strong coupling case, this will enable *the effective potential to adopt a negative value* when the field stabilizes at the minimum. The integral grows up indefinitely as the upper limit goes to infinity, *i.e.* it has an ultraviolet divergence. Because of the non-renormalizability of the theory, we cannot avoid this divergence, so we regularize by introducing a cut-off $\Lambda$. This parameter defines the energy scale below of which the theory is valid. We define the $x$ variable as

$$x \equiv \frac{m_\psi^2}{\Lambda^2} = \frac{m^2 g^2 \phi^2}{\Lambda^2}, \tag{7}$$

and the potential becomes

$$V_0 = \frac{\Lambda^2 x}{2g^2}, \tag{8}$$

$$V_1 = -\frac{\Lambda^4}{16\pi^2}\left[x + x^2 \log\left(\frac{x}{1+x}\right) + \log\left(1+x\right)\right]. \tag{9}$$

Notice that the one-loop potential $V_1$ is negative since it corresponds to the contribution of the original fermion field $\psi$, and we choose to parameterize it in terms of the effective scalar field $\phi$ c.f. Equation (4).

For the sake of concision we also define

$$A \equiv \frac{\Lambda^4}{16\pi^2}, \quad f\left(x\right) = x + x^2 \log\left(\frac{x}{1+x}\right) + \log\left(1+x\right). \tag{10}$$

In this way, taking quantum corrections into account we obtain an effective potential given by

$$V = V_0 + V_1 = \frac{\Lambda^2 x}{2g^2} - Af\left(x\right), \tag{11}$$

[1]Remember that the dimension of a scalar field equals that of mass.

[2]It will become clear that by doing so does not affect qualitatively the implied physical processes.

with the complete potential

$$V = V_0 + V_1 = \frac{\Lambda^2 x}{2g^2}\left(1 - \frac{g^2\Lambda^2}{8\pi^2}\right) - \frac{\Lambda^4}{16\pi^2}\left[x^2 \log\left(\frac{x}{1+x}\right) + \log(1+x)\right]. \tag{12}$$

As a function of $\phi$ it can be written explicitly as

$$V(\phi) = \frac{1}{2}m^2\phi^2 - \frac{\Lambda^4}{16\pi^2}\left[\left(\frac{mg\phi}{\Lambda}\right)^2 + \left(\frac{mg\phi}{\Lambda}\right)^4 \log\left(\frac{\left(\frac{mg\phi}{\Lambda}\right)^2}{1+\left(\frac{mg\phi}{\Lambda}\right)^2}\right) + \log\left(1+\left(\frac{mg\phi}{\Lambda}\right)^2\right)\right]. \tag{13}$$

Equation (13) gives the complete NJL scalar potential, and we are interested in studying its cosmological implications. Let us determine the asymptotic behaviour of the scalar potential $V$ in Equation (13). To analyze the potential we seek for extremum points. For the function $f(x)$ in Equation (10) we have the derivative

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = 2\left[1 + x\log\left(\frac{x}{1+x}\right)\right], \tag{14}$$

and for the derivative of $V$ we have

$$\frac{\partial V}{\partial\phi} = \frac{m^2\Lambda^2\phi}{4\pi^2}\left\{\frac{4\pi^2}{g^2\Lambda^2} - 1 - \left(\frac{mg\phi}{\Lambda}\right)^2 \log\left(\frac{\left(\frac{mg\phi}{\Lambda}\right)^2}{1+\left(\frac{mg\phi}{\Lambda}\right)^2}\right)\right\}, \tag{15}$$

$$\frac{\partial V}{\partial\phi} = \frac{m^2\Lambda^2\phi}{4\pi^2}\left\{\frac{4\pi^2}{g^2\Lambda^2} - 1 - x\log\left(\frac{x}{1+x}\right)\right\}. \tag{16}$$

The condition $\frac{\partial V}{\partial\phi} = 0$ implies the following equations:

$$\text{i)}\ \phi = 0, \quad \text{or} \quad \text{ii)}\ \frac{4\pi^2}{g^2\Lambda^2} - 1 = x\log\left(\frac{x}{1+x}\right). \tag{17}$$

The first one says that the origin $\phi = 0$ is an extremum, and if we take the second derivative $\frac{\partial^2 V}{\partial\phi^2}$

$$\left.\frac{\partial^2 V}{\partial\phi^2}\right|_{\phi=0} = \frac{m^2 g^2\Lambda^2}{4\pi^2}\left(\frac{4\pi^2}{g^2\Lambda^2} - 1\right), \tag{18}$$

we see that if $\frac{4\pi^2}{g^2\Lambda^2} > 1$ then the extremum at $\phi = 0$ corresponds to a minimum, while for $\frac{4\pi^2}{g^2\Lambda^2} < 1$ we have a maximum at the origin. The equation above suggest to define a critical value of the coupling $g_c$ as

$$g_c^2 \equiv \frac{4\pi^2}{\Lambda^2} \tag{19}$$

so that we see that for a weak coupling $g < g_c$ we have a minimum at the origin, while at strong coupling $g > g_c$ we have a maximum. The type of extrema at the origin of the potential corresponds to the value of the coupling.

Now let us determine the second (possible) extreme of the potential. Since the r.h.s of the second equation in Equation (17) is negative (*i.e.* $x\log\left(\frac{x}{1+x}\right) \le 0$) this equation has a solution only for a strong coupling $g > g_c$. A value for $x$ (or that of the scalar field $\phi$), at the minimum cannot be solved analytically, since the second equation in Equation (17) is a transcendental equation. One way to determine a solution is to seek for the intersection between the curve of the function $x\log\left(\frac{x}{1+x}\right)$ r.h.s. in the second Equation (17), and the constant in the l.h.s. In this case do exist an intersection (only one, as the r.h.s. is a monotonic function), giving a solution for the $x$ variable, leading in its turn to a non-trivial solution in $\phi = \phi_{\min}$ which is a *minimum*.[3] The extremum in this case corresponds to a minimum. Notice that in all cases we have at large $x$ the limit $V \to \infty$ for $x \to \infty$ regardless of the value of the coupling $g$.

Therefore, we have: if $g < g_c$, the potential minimizes in the origin $\phi = 0$; whereas for $g > g_c$, the potential minimizes in a non trivial value $\phi = \phi_{\min}$. The value of the coupling $g = g_c$ define a critical value separating between both behaviours of the potential (in Figure 2 we show all the three cases $g < g_c, g = g_c, g > g_c$). When for $g > g_c$ we see that the full potential $V = V_0 + V_1$ becomes negative, due to the contribution of $V_1$, and a fermion condensate $\bar{\psi}\psi \neq 0$ is formed and is parameterized by the scalar field $\phi = (g/m)\bar{\psi}\psi$, c.f. Equation (4).

To estimate the value of the potential at the minimum for $g > g_c$, the equation ii) in Equation (17) should be solved. However, since it is a transcendental equation in the variable $x$, an algebraic expression cannot be written, and we need to use numerical

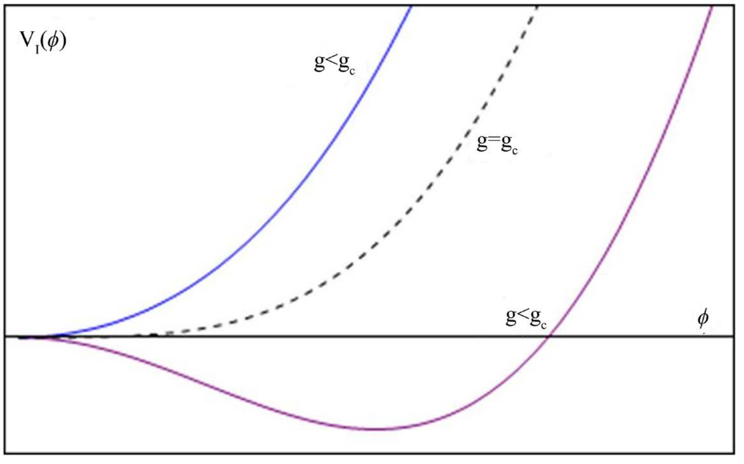



Figure 2. Effective potential (13) as a function of $\phi$. The critical value of the coupling $g_c$, separates two kinds of behaviours.

[3]According to definition Equation (7), $x$ is a quadratic function in $\phi$: $x \sim \phi^2$, so for a given value of $x$ we have two solutions in $\phi$ related by a change of sign. Due to this symmetry, we will allow ourselves to refer to only one solution.

procedures. Let us introduce a parameter $s$ to write $g$ in the form

$$g = s \cdot g_c = s \cdot \frac{2\pi}{\Lambda}. \tag{20}$$

In this way we make sure to have a strong coupling by taking $s > 1$. Now, for a given value of the coupling with $g > g_c$, there exists a definite value of *x*, say $x_0$, satisfying the condition ii) in Equation (17), which is the solution for the minimum. Therefore, the potential valuated at this point yields the minimized potential, *i.e.* $V_{\min} = V(x_0)$. Then, by substituting ii) Equation (17), and using Equation (20) in the expression for the potential Equation (12), we can write $V_{\min}$ in the suitable form

$$V_{\min} = \frac{\Lambda^4}{16\pi^2}\left[\frac{x_0}{s^2} - \log(1 + x_0)\right], \quad (g = s \cdot g_c, s > 1), \tag{21}$$

which provides a good idea of how $V_{\min}$ is related to the energy scale.

From Equation (4) the field $\phi \sim \bar{\psi}\psi$, is a Lorentz invariant quantity, so $\phi$ is scalar field. When the field $\phi$ is stabilized, a non trivial expectation value reflects the presence of a fermion condensate.

Now, if the field has an expectation value $\langle\phi\rangle = 0$, it means that the state of paired fermions $\bar{\psi}\psi$ is not present, so we have a system consisting in the original massless fermion particles with a 4-Fermi interaction, and a condensate is not energetically favoured. This happens for a "weak" coupling $g < g_c$. On the other hand, if the expectation value $\langle\phi\rangle \neq 0$, then we have a fermion condensate represented effectively by the scalar field. This happens for a "strong" coupling $g > g_c$, and a fermion condensate is dynamically formed since it reduces the energy of the system.

Thus, we see that two different fluid phases (massless fermions or fermion condensate) are obtained depending on the strength of the coupling. Next, we investigate the cosmological dynamics of each of these fluids.

## 3. Standard Cosmology

The widely accepted current standard cosmological model (the Big Bang theory) is based in Einstein's theory of General Relativity. If conditions of spatial homogeneity and isotropy are assumed, the space-time metric adopt the well-known simple form

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a(t)^2\left[\frac{\mathrm{d}r^2}{1 - kr^2} + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2\right] \tag{22}$$

where the variables $r, \theta, \phi$ are comoving coordinates parameterizing the spatial section of space-time, and $k$ takes the values +1, 0, −1 for spaces of constant positive (spherical), zero (flat), or negative (hyperbolic) curvature. When this metric is used in the Einstein's equations, the so called FRWL equations (Friedmann-Robertson-Walker-Lemaitre) can be obtained. As these assumptions agree with observations[4] to a very high precision, we will use this same theoretical framework. Because the necessary equations are well known and their deduction can be found in standard text books, in the

[4]CMBR is a smooth bath of radiation, whereas Large Scale Structure reveal uniform distribution of matter at cosmological scales, with $\gtrsim 100$ Mpc.

following we limit ourselves to write them and to give only a brief explanation.

The equation

$$\frac{\dot{a}^2}{a^2}+\frac{k}{a^2}=\frac{8\pi G}{3}\rho, \tag{23}$$

relates the expansion rate (in time) of the scale factor *a*, and the curvature *k* of the universe, to the total energy density $\rho$. Along this paper we will *always* take a flat geometry $k=0$, as suggested on the one hand from the *theory* of early cosmological inflation, and on the other hand (and most important) from *observation* of the CMBR.

Introducing the usual definition relating the Hubble parameter $H$ with the rate of change in time of the scale factor $a$

$$\dot{a}=a\cdot H, \tag{24}$$

Equation (23) (with $k=0$, flat universe) becomes

$$H^2=\frac{8\pi G}{3}\rho. \tag{25}$$

The continuity equation for a fluid with energy density $\rho$ and pressure $P$ is

$$\dot{\rho}+3H\left(\rho+P\right)=0. \tag{26}$$

For a perfect fluid "$\alpha$" satisfying a *barotropic* equation of state $P_\alpha=w_\alpha\rho_\alpha$, with $w_\alpha$ a constant, Equation (26) can be solved analytically. We sometimes will refer to such a fluid with the name of "barotropic fluid". From the cosmological point of view, the substances contained in the universe can be described as radiation, which has $w_r=1/3$, and matter (dust) having $w_m=0$ (besides the Dark Energy component). For those we have respectively

$$\rho_r=\rho_{ri}\left(\frac{a}{a_i}\right)^{-4}, \quad \rho_m=\rho_{mi}\left(\frac{a}{a_i}\right)^{-3}. \tag{27}$$

A scalar field $\phi$, with a self-interaction potential $V\left(\phi\right)$, has energy density $\rho_\phi$ and pressure $P_\phi$ given by

$$\rho_\phi=E_k+V\left(\phi\right), \quad P_\phi=E_k-V\left(\phi\right), \quad E_k=\frac{1}{2}\dot{\phi}^2 \tag{28}$$

where we have also defined the kinetic energy $E_k$ in the third equation. Considering an universe containing radiation, matter and a scalar field, the total energy density is written

$$\rho=\rho_r+\rho_m+\rho_\varphi. \tag{29}$$

For a given component fluid "$\alpha$", it is useful to know its *relative density*, defined as the ratio of its energy density to the total energy density:

$$\Omega_\alpha=\frac{\rho_\alpha}{\rho}=\frac{8\pi G\rho_\alpha}{3H^2}, \tag{30}$$

where we have used Equation (25) in the second equality. In a flat universe one has the condition

$$\Omega_r+\Omega_m+\Omega_\phi=1. \tag{31}$$

It is interesting to note that while Equation (31) remains valid even when we have a negative $\rho_\alpha$, the quantity $\Omega_\alpha$ is no longer constrained to the values $0 \le \Omega_\alpha \le 1$. In the work presented here, the fluids can have a negative energy density, giving $\Omega_\alpha < 0$, or a total energy density $\rho$ that vanish at finite values of the scale factor $a(t)$, in which case we would have $\Omega_\alpha \to \pm\infty$.

Taking the time derivative in Equation (25) and using Equation (26), it can be found

$$\dot{H} = -\frac{1}{2}\left(\rho_m + \frac{4}{3}\rho_r + \dot{\phi}^2\right). \tag{32}$$

Note that the r.h.s. in Equation (32) is always negative. The equation of motion for a spatially homogeneous scalar field, (a modified Klein-Gordon equation) is given by

$$\ddot{\phi} + 3H\dot{\phi} + \partial_\phi V = 0. \tag{33}$$

It is also useful an equation for the acceleration of the scale factor:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3P\right). \tag{34}$$

Differential Equations (24), (32), (33), together with (27) constitute a complete set which can be solved numerically (since we cannot always write an analytical solution). Nevertheless, it is convenient to attempt to outline the general behaviour of the dynamical system. Thus, before going to solve for our NJL potential, let us point out the following generic facts:

The evolution of the scalar field is such that it will minimize the scalar potential $V(\phi)$, so for an arbitrary initial value $\phi_i$, the field will roll to lower values of the potential, in such a way that eventually it will adopt a constant value ($\phi = \phi_{\min}$ being the minimum). Given than the scale factor is a positive defined quantity, the energy densities for matter and radiation Equation (27) are always positive quantities and never equal to zero for finite values of the scale factor $a(t)$. So, the total energy density Equation (29) remains always positive as long as the condition

$$\rho = \rho_r + \rho_m + \rho_\phi > 0 \tag{35}$$

is satisfied. Thus, Equation (25) says that $H = 0$, that is $\dot{a} = 0$, *never* happens (Equation (24)) as long as $\rho \ne 0$. This implicates that $\dot{a} > 0$ *always*. This means that the scale factor $a(t)$ never reaches an extremum value along its time evolution (taking an initial condition $H_i > 0$, since we know that the universe is expanding at present time).

Nevertheless, it is interesting o note that there is no known physical principle forbidding the existence of a fluid with a negative potential $V(\phi) < 0$, at least for some values of the field $\phi$. In this case, it could well happen that Equation (35) become an equality, meaning $\rho = 0$ for finite values of $a(t)$, which in turn implies $H = 0$, and $\dot{a} = 0$; *i.e.*, the scale factor reaches an extremum value (indeed a maximum, since as seen before, it was initial growing). Now, Equation (32) imposes an always decreasing Hubble parameter $H$ (because the right hand side is always negative), so that after being $H = 0$ it must be $H < 0$, and therefore $\dot{a} < 0$, *i.e.* the scale factor decrease. In other words, the universe must be contracting after reaching its maximum size. Observe that this result is a consequence only of the negativity of the potential, and it is independent of its specific form. This collapsing universe is valid even for a flat universe $k = 0$. To conclude, if a fermion condensate is energetically favored then the minimum of potential $V(\phi)$ is negative and the universe will recolapse.

## 4. Dynamics of Massless Fermions Phase (Weak Coupling $g < g_c$)

As we have seen in Section 2, for a weak coupling $g < g_c$ the minimum of the potential $V(\phi_{\min}) = 0$ is located at the origin with $\phi_{\min} = 0$, and $V$ does not take negative values. Therefore, the total energy density and $H$ never vanish for finite values of the scale factor *a*, and we have $\dot{a} > 0$ due to Equation (24). So the scale factor $a(t)$ is always growing, going to an infinite size in an infinite time. Now, from Equation (34), it can be seen that, in order to have $\ddot{a} < 0$, *i.e.* the universe to slow down its expansion rate, then

$$\frac{2}{3}\rho_r + \frac{1}{2}\rho_m + \dot{\phi}^2 > V(\phi) \tag{36}$$

is a condition to be satisfied. This, of course, in not always the case: we could take an initial field amplitude $\phi_i$ as big to make the initial value of the potential $V_i = V(\phi_i)$ big enough so that inequality (36) does not hold, and we would have instead $V > \frac{2}{3}\rho_r + \frac{1}{2}\rho_m + \dot{\phi}^2$. In this case we could have an acceleration of the scale factor, *i.e.* an accelerating universe, though it would be an "early" acceleration, as it would be present an initial times, *i.e.* before letting the fluid densities to dilute and field to evolve. As time passes, the field rolls down minimizing the potential, and eventually acquires some value $\phi < \phi_i$ such that condition (36) becomes fulfilled.

Given that the densities of matter and radiation never reach a null value in a finite time, and that the field amplitude tends to be stabilized around the minimum (*i.e.* $\phi \to 0$), for a big enough amount of time, we expect a vanishing potential and velocity, $V \sim 0$, $\dot{\phi} \sim 0$ to be a good approximation to a final situation, in which (36) is still satisfied.

We show an example of numerical solution in the figures. In **Figure 3** we see that the field has a damped oscillation around $\phi = 0$, and in consistency with this, its kinetic energy (velocity) diminish in time and we show in **Figure 4** the evolution of the relative densities $\Omega_{rad}, \Omega_{mat}, \Omega_\phi$ for radiation, matter and $\phi$. Simultaneously, the potential valuated at $\phi \sim 0$ goes to lower values (according to $V(\phi_{\min}) = 0$). We can see that although the universe is expanding, it always ends up in a non-accelerating regime (**Figure 5**). A Taylor expansion for the potential about $\phi = 0$ gives

$$V \simeq \frac{1}{2}m^2\left(1 - \frac{\Lambda^2 g^2}{4\pi^2}\right)\phi^2, \tag{37}$$

where the whole coefficient multiplying on $\phi^2$, is a positive quantity, as $g^2 < 4\pi^2/\Lambda^2$. The coefficient of state $\omega_\phi$ defined below Equation (26), for the field $\phi$, written explicitly is

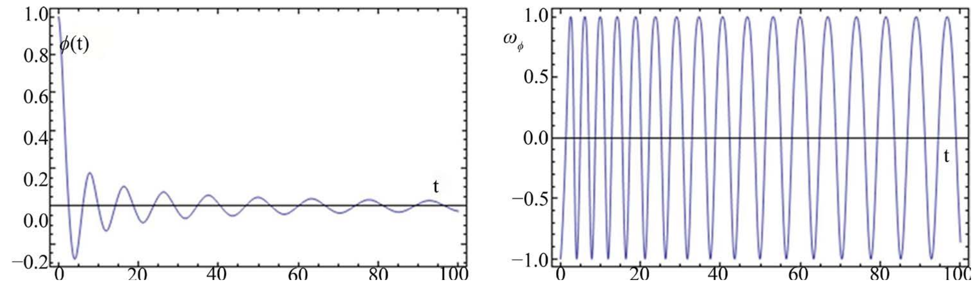



**Figure 3.** *Left*: Scalar field amplitude $\phi$. *Right*: State equation coefficient $\omega_\phi$. Both variables are shown as functions of time.

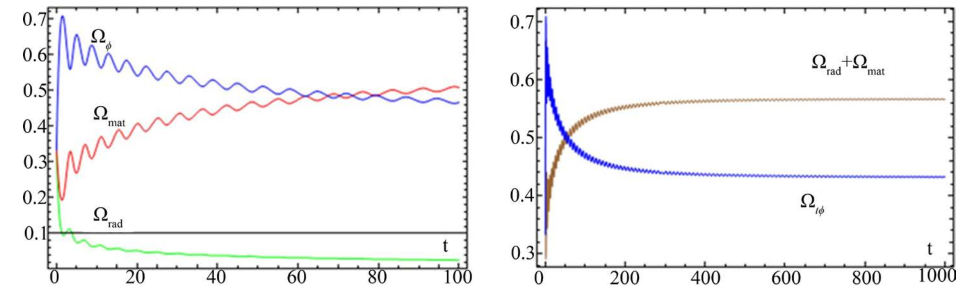



**Figure 4.** *Left*: Relative densities $\Omega_\alpha$ for radiation, matter and $\phi$. *Right*: Total relative density for barotropic fluids (matter and radiation) and for the field $\phi$. The horizontal axis in both graphics represents time. Note that we show a different scale of time in each plot for the same solution.

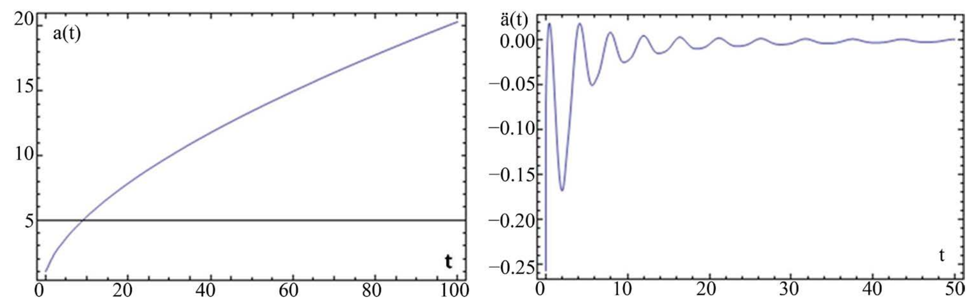



**Figure 5.** *Left*: Scale factor $a(t)$. *Right*: Acceleration of the scale factor $\ddot{a}(t)$. Both variables are shown as functions of time. Note that $\ddot{a}(t)$ adopt mostly negative values (tends to zero from below).

$$\omega_\phi = \frac{P_\phi}{\rho_\phi} = \frac{E_k - V}{E_k + V}. \tag{38}$$

Since at late times, when the field oscillates around its minimum with a quadratic potential, the average value is $\langle \omega_\phi \rangle = 0$ and $\rho_\phi$ evolves as matter with $\rho_\phi \propto a^{-3}$ [24].

Within the context of Early Cosmic Inflation theory, the so called *Slow Roll* parameters are defined as follows:

$$\epsilon = \frac{M_p^2}{2}\left(\frac{V'}{V}\right)^2, \quad \eta = M_p^2\left(\frac{V''}{V}\right), \tag{39}$$

which have to satisfy the conditions $\epsilon < 1$, $|\eta| < 1$ in order to the potential may cause a positive acceleration. Even though they are valid for a single field, without additional fluids (matter and/or radiation), we show them in Figure 6 the Slow Roll parameters, for the seek of completeness.

## 5. Fermions Condensate Dynamics (Strong Coupling, $g > g_c$)

The strong coupling case leads to a fermion condensate and therefore to a negative potential $V$ at its minimum. The potential has at the origin $V(\phi = 0) = 0$ and decreases to negative values for $0 < \phi < \phi_{\min}$. For $\phi > \phi_{\min}$ it grows monotonically, eventually passing from negative to positive values. Let us consider at first the simpler approach of a universe containing only a scalar field ( $\rho_r = \rho_m = 0$, *i.e.* no additional fluids), evolving under a *generic* potential possessing a negative value $V_{\min} = V(\phi_{\min}) < 0$ when minimized. If the initial velocity $\dot{\phi}_i = 0$, then the kinetic energy of the field has a null value as well, so we have for the initial energy density $\rho_{\phi_i} = V_i$. The initial amplitude for the field $\phi_i$ cannot be such that makes $V(\phi_i) < 0$, because it would lead to an imaginary value for *H*, according to Equation (25). Thus, we must take always $\phi_i$ such that $V_i > 0$. As before we begin with $H_i = 1 > 0$, therefore Equation (24) says that $a(t)$ initially is increasing in time. The equation (32) is written $\dot{H} = -(1/2)\dot{\phi}^2$, so that *H always diminish* in time. As the potential is minimized, it goes from positive to negative values, and from Equation (25) eventually it will be $H = 0$, and after this $H < 0$, corresponding respectively to $\dot{a} = 0$ and $\dot{a} < 0$. In words this means that after an initial period of expansion (increasing scale factor), a maximum value is reached, followed by a period of contraction. Since $\dot{H}$ remains always negative, then $a(t)$ will continue decreasing, so that it necessarily will collapse. In other words, it will be $a = 0$ in a *finite* time in the *future* (because the evolution is forward in time: the field minimizes, not otherwise).

Now, while the expanding phase is taking place, the field is rolling down, eventually entering in a damped oscillatory regime nearly the minimum, where the potential has become negative, $V_{\min} < 0$. Because of the damping, the kinetic energy tends to a zero value, $E_k \to 0$. Thus, the energy density of the field $\rho_\phi = E_k + V$ goes from positive values (near $\phi_i$) to negative values (near $\phi_{\min}$), so at some time in between, it is

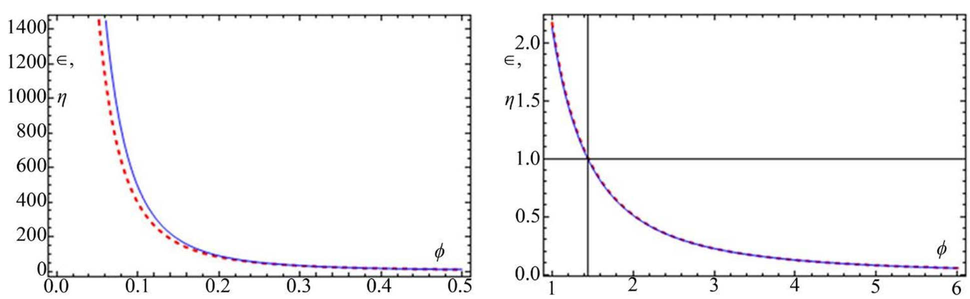



Figure 6. Slow roll parameters $\epsilon$ (dashed-red curve), and $\eta$ (continuous-blue). *Left*: From $\phi = 0$ to $\phi = 0.5$. *Right*: From $\phi = 1$ to $\phi = 6$. Only in the region $\phi \gtrsim 1.4$ approx. (and further on) one can expect the acceleration conditions $\epsilon < 1$, $\eta < 1$ to be satisfied.

$\rho_\phi = 0$. The total energy density, as well as the individual densities for each fluid (if there were additional fluids), would go to *diminish* in time (as can be seen for radiation and matter in Equation (27) with $\rho_\alpha \sim a^{-n}$, and $a(t)$ increasing). By a similar reasoning, because $a(t)$ is decreasing in the *contracting* phase, the energy densities behave the opposite way, *i.e.* they all *increase* in time. Therefore, we expect $\rho_\phi = 0$ to happen twice. In its turn, this implicate that the coefficient of state $\omega_\phi$, Equation (38) become a divergent quantity also twice, around this two points, and near them, $\omega_\phi$ is not anymore a useful parameter to characterize the fluid represented by the field $\phi$. Below we show a numerical solution example (Figures 7-11).

As we mention before, in Section 3, a similar circumstance arises in dealing with the relative densities $\Omega_\alpha$: it is considered that in order to this parameter to make sense, a relative density should adopt values $0 \le \Omega_\alpha \le 1$. However, as can be seen in Equation (30), if at some time is $H = 0$, then nearly this value, each $\Omega_\alpha$ turns into a divergent variable. The situation is even weirdest for the field, because near the minimum it is $\rho_\phi \sim V_{\min} < 0$, the energy density of the field is similar to the potential, which is negative. This would make $\Omega_\alpha \to -\infty$ (a divergent and *negative* relative density!).

Consider now a universe containing matter and radiation in addition to our NJL fluid. An interesting question is, may the presence of these fluids prevent the universe to collapse? Remember that the condition for an increasing scale factor can be reduced to the inequality (35). If the scale factor is supposed to grow forever, this condition must be hold *always*. Now, according to the explanations given above, initially the scale factor is growing indeed. Thus, from Equation (27) we see that the densities of both barotropic fluids (matter and radiation) must be decreasing. At the same time, because the field is stabilizing in the minimum of the potential, the kinetic energy of the field $E_k = (1/2)\dot{\phi}^2$ is diminishing to zero, whereas the potential is going to a constant value $V \to V_{\min}$, in such a way that necessarily, condition (35) ceases to hold. Therefore, even in presence of additional barotropic fluids (does not matter the relative amount with respect to that of the fluid associated with the field), the collapsing universe situation cannot be avoided.

The previous qualitative generic analysis is verified by the numerical solution for our NJL potential in particular (Figures 7-11). By observing the graphics, we found an unpredicted interesting non-trivial behaviour of the field amplitude: while the scale factor undergoes the expanding, and contracting phases successively, an damped oscillating phase around $\phi_{\min}$ is taking place, as expected. But then, at some point in the contracting phase, the field amplitude goes to bigger values, and as the scale factor approach to $a = 0$, the field is taken out from the minimum and it begins to increase monotonically![5] Is this an acceptable result? Intuitively, as *a* is decreasing, it is reasonable to expect all densities to be growing. In particular, if the field density $\rho_\phi = E_k + V$ is getting bigger, it should be due to an increase in the field velocity (so $E_k$ gets bigger), or in the field amplitude (so $V$ gets bigger); or both. This behaviour can indeed be

[5]It could decrease instead, depending on the initial conditions. Whatever the case, the monotonic growing in absolute value is an unexpected behaviour, which do happen indeed.

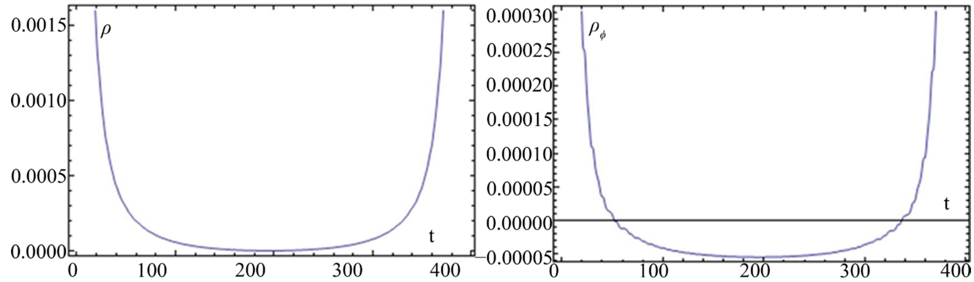



**Figure 7.** *Left*: Total energy density $\rho = \rho_r + \rho_m + \rho_\phi$. It is a positive quantity, but vanishes at a single point, near $t \simeq 200$ approx. *Right*: Energy density of the field. It is a null quantity ( $\rho_\phi = 0$ ) twice: one time in the expansion phase (near $t = 60$ approx.), and again in the contraction phase (about $t = 340$ approx.); and becomes a negative quantity in between.

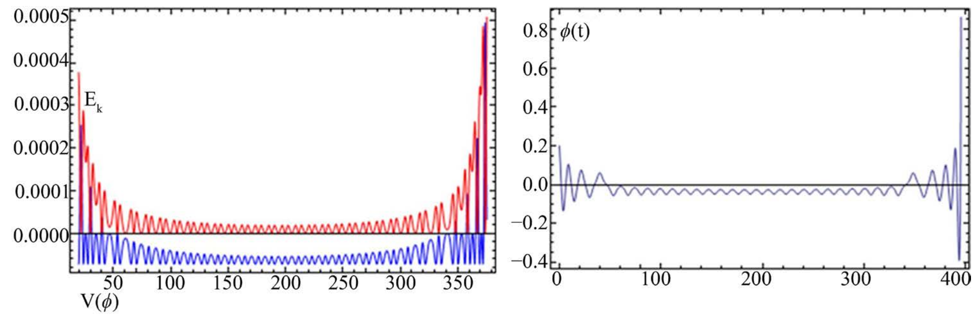



**Figure 8.** *Left*: Although the kinetic energy (red-upper curve) is zero initially, it overtakes the potential energy (blue-lower curve) and remains dominant all the way even to the collapsing time when $a(t) = 0$. *Right*: The field oscillates around $\phi_{\min}$ and is becoming divergent as getting close to $t \simeq 400$, which is the time when $a(t) \to 0$.

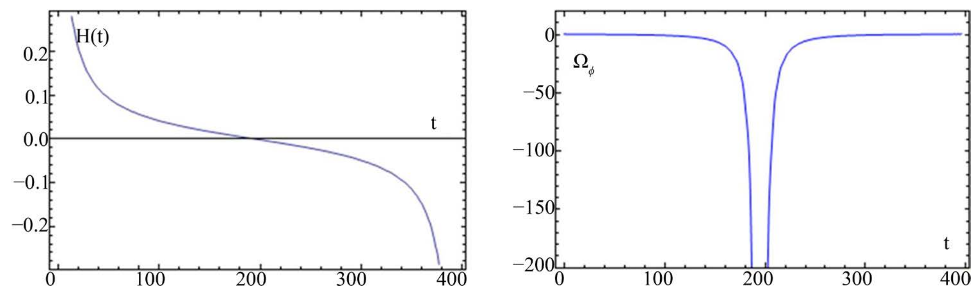



**Figure 9.** *Left*: Hubble parameter. It is a null quantity about $t \simeq 200$ approx. *Right*: Relative density of the field. As $H(t)$ vanish, $\Omega_\phi$ becomes a divergent quantity near the null point.

explained observing Equation (26). The energy evolution of a barotropic fluid $\rho_b$ is given by

$$\dot{\rho}_b = -3H\left(\rho_b + P_b\right) = -3H\rho_b\left(1+w\right), \tag{40}$$

and for a scalar field with energy density $\rho_\phi = E_k + V(\phi)$ and pressure $P_\phi = E_k - V(\phi)$ and $E_k = \dot{\phi}^2/2$

$$\dot{\rho}_\phi = -3H\left(\rho_\phi + P\right) = -3H\dot{\phi}^2 = -6HE_k. \tag{41}$$

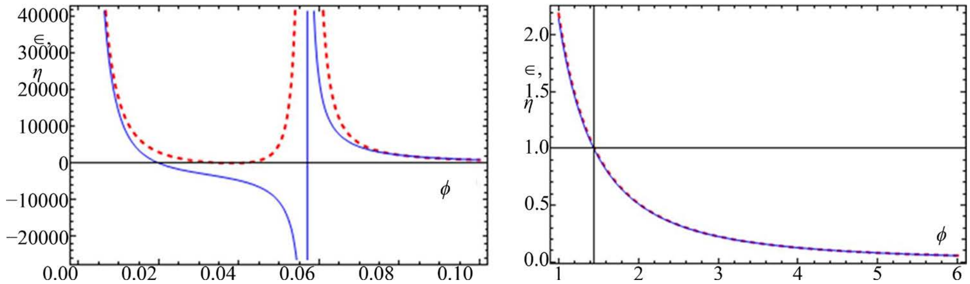



**Figure 10.** Slow roll parameters $\epsilon$ (dashes-red curve), and $\eta$ (continuous-blue). *Left*: From 0 to 0.1 in $\phi$. *Right*: From 1 to 6 in $\phi$. Only in the region $\phi \gtrsim 1.4$ approx. (and further on) one can expect the acceleration conditions $\epsilon \ll 1$, $\eta \ll 1$ to be satisfied.

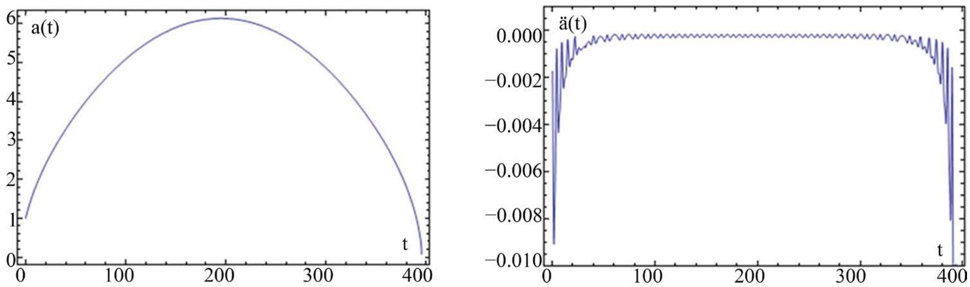



**Figure 11.** *Left*: Scale factor. *Right*: Acceleration $\ddot{a}(t)$. Both plots are to be interpreted as describing a universe which expands without acceleration (note that $\ddot{a}(t)$ is never greater than zero), reaching a maximum value about $t \simeq 200$ approx., thereafter falling in a contracting phase all the way long to collapse.

We can see from Equations (40) and (41) that for a positive barotropic fluid $\rho_b$ with an EQS $|w| < 1$, the sign of $\dot{\rho}_b$ and $\dot{\rho}_\phi$ are negative as long as $H$ is positive while they become negative for $H < 0$. Therefore $\rho_b, \rho_\phi$ are decreasing functions as a function of time for $H > 0$ and increasing for $H < 0$. Since we have seen that $\dot{H}$ is negative, this implies that $H$ is always a decreasing function of time. If $H$ can vanish at a finite time only if $\rho_\phi$ becomes negative, *i.e.* if the potential $V$ becomes negative and $E_k = -V$ at say $t = t_c$. After this time $H(t > t_c)$ becomes negative and will remain negative for $t > t_c$ and $\rho_b$ and $\rho_\phi$ will start growing with time for $t \geq t_c$.

Figure 8 show both kinetic and potential energies, and we can see that even though the initial kinetic energy is zero, it overtakes the potential energy and remains so until the collapsing moment $t_{final}$ when $a(t) = 0$ at late times. Nevertheless, the potential energy also grows as the time is approaching $t_{final}$, so the field amplitude is eventually expelled from oscillating about the minimum.

## 6. NJL Fluid with a Cosmological Constant

Due to its theoretical properties and observational requirements, a Cosmological Constant is a very usual and useful ingredient included in cosmological models, and it is worth to consider such contribution in our model. Its defining property is an energy density $\rho_\Lambda$ which does not vary in time, and a coefficient of state $\omega_\Lambda = -1$, which gives a pressure $P_\Lambda = -\rho_\Lambda$. In a universe containing *only* a Cosmological Constant, the equation (34) is written $\ddot{a} = (8\pi G/3) a \times \rho_\Lambda$ which, as $\rho_\Lambda > 0$, implicates $\ddot{a}(t) > 0$ *always*. Therefore, such an universe is always accelerating its expansion. In fact, in this case the Equation (24) may be solved analytically, after substituting Equation (25), giving the well known solution $a(t) = a_i \exp\left(t\sqrt{8\pi G \rho_\Lambda/3}\right)$. How do the presence of a Cosmological Constant affect our previous considerations of a universe including our NJL fluid, besides matter and radiation components? Will the universe accelerate or collapse, even in the presence of a scalar field with a negative potential $V < 0$? Because the density $\rho_\Lambda$ is constant, we have that the differential equations are not modified, other than just adding a term in the expression for *H*, equation (25). In particular, the equation of motion Equation (33) remains unchanged, so the field dynamics is not affected. As before, we have to deal with two cases.

**a) Free Fermions ($g < g_c$).** As studied before, the potential is $V \geq 0$, and its minimum value is $V_{\min} = 0$. Also, with the pass of time, both matter and radiation densities dilute, going to vanish. From Equation (34), it can be deduced the condition for universe to decelerate:

$$\rho_\Lambda < \rho_r + \frac{1}{2}\rho_m + 2E_k - V(\phi) \quad (\text{for } \ddot{a} < 0). \tag{42}$$

Given that the left hand side in this inequality is diminishing in time, whereas the right hand side remains constant, we have that eventually this inequality cannot hold anymore, and becomes an equality, meaning $\ddot{a} = 0$. This points the beginning of the acceleration period, *i.e.* $\ddot{a} > 0$, where the inequality (42) gets inverted. Had the initial conditions been such that inequality (42) were the opposite, then there would be always an acceleration holding always, because the LHS would never go back to grow.

Thus, we see that for a free fermions NJL fluid with a Cosmological Constant, the universe necessarily accelerate, the precise moment depending on the amount of energy densities $\rho_m$, $\rho_r$, with respect to that of $\rho_\Lambda$. This can be specified in the initial conditions, which in their turn can be chosen to solve for a realistic model fitting the observations.

**b) Fermion Condensate ($g > g_c$).** We found before that, for a strong coupling, the potential is negative when minimized, $V_{\min} < 0$. Do the universe necessarily accelerate also in this case? In order for this to happen, condition (42) eventually must turn into an equality, meaning $\ddot{a} = 0$. This is a minimal condition to be satisfied, because it points at least the beginning of an acceleration; it remains to be sure that acceleration will be sustained. Let us label all quantities with a subindex "*ac*" at time $t_{ac}$, when $\ddot{a} = 0$ (vgr. $V(t_{ac}) = V_{ac}$). From Equation (36), we have[6]

$$\rho_\Lambda \geq \rho_{rac} + \frac{1}{2}\rho_{mac} + 2E_{kac} - V_{ac} \quad (\text{for } \ddot{a} \geq 0). \tag{43}$$

Remember that the potential take positive values as well as negative ones, so both possibilities must be taken into account. Certainly one can find such set of values of $V$

[6] $\rho_\Lambda$ does not need a label because it is a constant.

for a given $\rho_\Lambda$ to satisfy the inequality. However, if we rather want to consider realistic models, we should consider plausible values from observations (besides, we would not like to complicate our lives by considering unrealistic generic situations).

From definitions (30) it can be found that $\Omega_r/\Omega_m = (1+z)r$, where $z$ is the redshift, and $r = \Omega_{r0}/\Omega_{m0}$ says the amount of radiation with respect to that of matter. The subindex "0" refers to current values, *i.e.* quantities measured "today". Now, the estimate for $z$ (the time when acceleration begins) is around $z \sim 1$; and it has been measured $r \sim 10^{-4}$ (for the seek of simplicity, here we are interested only in orders of magnitude). Then we have $\Omega_{rac} \sim 10^{-4} \times \Omega_{mac} \ll \Omega_{mac}$, or $\rho_{rac} \ll \rho_{mac}$. Now, remember that a decelerating period dominated by matter is supposed to have taken place before $\ddot{a} = 0$. In order for this to happen, condition (42) should have to be true before condition (43). For $z \sim 1$ (it could be even as big as, let's say $z \sim 10$, as this would not change the essence of the argument) and using condition (42) we would have

$$\rho_m > 2\left(\rho_\Lambda + V - 2E_k\right) \quad \left(\text{in order to be } \ddot{a} < 0\right). \tag{44}$$

If a positive acceleration eventually come up, the above expression is expected to become an equality. Now, suppose $V > 0$. Then, unless $E_k$ decrease even fast, the RHS in the inequality should be decreasing as time passes, because the potential is minimizing. But $E_k$ cannot behave like that indeed, as the field is under a damped rolling, not to mention that $E_k$ is never a negative quantity, so the sum of terms $V - 2E_k$ will end up decreasing (would the values of these terms been such that the equality *somehow* would be accomplished at some time, in this case the acceleration *could not be attached* to $\rho_\Lambda$ anyway). On the other hand, for $V < 0$, the inequality would become even more strong in time, because again, the potential is minimizing: $V \to V_{\min}$, and $0 > V > V_{\min}$. Therefore, if initially the inequality (44) begins being satisfied, it will remain being so always; in other words, the universe will never accelerate.

What about a collapse *in the future*? May the presence of a cosmological constant prevent a decreasing scale factor (time going forward)? For a growing scale factor we have $\dot{a} > 0$, which is true indeed because we take $H_i > 0$ is the initial value of $H$.[7] As we explained before, if the scale factor is to reach a maximum $a = a_{\max}$, it must be $\dot{a} = 0$. Let us name $t_{am}$ the time when this is accomplished (if so), and label with a subindex "*am*" the variables valuated at this time. We have for the total energy density $\rho_{am} = 0$, thus $\rho_\Lambda + \rho_{ram} + \rho_{mam} + E_{kam} + V_{am} = 0$. The only way in which this could happen is for $V_{am} < 0$. In that case $V_{am} = -|V_{am}|$, so the equation, as a condition to be satisfied by $\rho_\Lambda$, can be written in the more intelligible form

$$\rho_\Lambda = |V_{am}| - E_{kam} - \rho_{mam} - \rho_{ram} \quad \left(\text{to get } \dot{a} = 0\right). \tag{45}$$

If we want to keep our analysis as simple as possible, we may ignore the contribution from radiation, $\rho_{ram} = 0$ (observe that, had an acceleration would be possible, then we should assume $t_{ac} < t_{am}$, *i.e.* acceleration before receding, otherwise the model would not be useful. So, if $\rho_{rac} \ll 1$ the approximation $\rho_{ram} \sim 0$ is even better, as $\rho_{ram} < \rho_{rac}$).

[7]Observe this initial condition must be taken to be positive, because otherwise, the universe would be *already* contracting.

Now, nothing forbids to exist a potential sufficiently deep $V_{\min} < 0$, so that the equality (45) can be accomplished. The exact time at which this is achieved will depend on the relative amounts $E_{kam}$, $\rho_{mam}$, with respect to $\rho_\Lambda$, *i.e.* on the initial conditions. However, we can estimate a limit value by making $\rho_{mam} \to 0$, $E_{kam} \to 0$, and a stabilized potential $V \to V_{\min}$. Then we have

$$\rho_\Lambda = |V_{min}| \quad (\text{Maximum allowed value for the universe to collapse}). \tag{46}$$

After $\dot{a} = 0$, *i.e.* $H = 0$ (Equation (24)), the universe must enter into a contraction phase because $H$ is *always* decreasing (Equation (32)), meaning $H_{am} \to H < 0$, *i.e.* $\dot{a} < 0$. So, eventually the universe will collapse in the future in a finite lapse of time. For $\rho_\Lambda > |V_{\min}|$, the scale factor would never go to contract, as in this case the total energy density $\rho$ would never vanish.

It is interesting to observe that a Cosmological Constant may be seen as a particular case of a scalar field evolving under a potential stabilized with a positive minimum. As we have seen, the NJL model has two different behaviours depending on the value of the coupling constant $g$. For weak coupling $g < g_c$ the potential $V(\phi)$ has a minimum at the origin with $V(\phi = 0) = 0$ and $V(\phi) \geq 0$ otherwise. On the other hand, at strong coupling $g > g_c$ one has a negative minimum $V(\phi)\big|_{\min} < 0$. So let us approximate the potential $V$ around the minimum and take the ansatz

$$V(\phi(t)) = V_o + \frac{1}{2} m^2 (\phi(t) - \phi_o)^2, \tag{47}$$

with $V_o$ a constant value (it would be $V_o = 0$ at weak coupling and $V_o < 0$ at strong coupling) and $\phi_o$ a constant. We can now ask ourself if we can have an accelerating universe. The evolution of the scalar field is just $\ddot{\phi}' + 3H\dot{\phi}' + m^2\phi' = 0$, with $\phi' \equiv \phi - \phi_o$ and we could redefine $\rho_\Lambda + \rho_\phi = \rho_\Lambda + E_k + V = \rho_\Lambda + V_o + E_k + \frac{1}{2} m^2 (\phi - \phi_o)^2 = \rho_\Lambda + V_o + E_k + \frac{1}{2} m^2 \phi'^2$ which corresponds to a massive scalar field with energy density $\rho_{\phi'} = E_k + \frac{1}{2} m^2 \phi'^2$ in the presence of a cosmological constant $\rho'_\Lambda = \rho_\Lambda + V_o$. A massive scalar field may accelerate the universe only at large values of $\phi'$ (larger than the Planck mass) when the Slow Roll parameters $\epsilon$ and $\eta$ are smaller than one, while at a late time when the scalar field oscillates around the minimum the energy density $\rho_{\phi'}$ redshifts as matter, *i.e.* $\rho_{\phi'} \propto 1/a^3$. In order to have $\ddot{a} > 0$ we must have the quantity $\xi \equiv \rho + 3p < 0$. So for a scalar field (with potential given in Equation (47)) a barotropic fluid, which we now take for simplicity as matter (without lose of generality), and a cosmological constant $\rho_\Lambda$, we have $\xi = \rho_m + 4E_k - 2(\rho_\Lambda + V)$. Since the potential $V_o$ vanishes at weak coupling and is negative at strong coupling, there is a cancelation between the two cosmological constants $\rho_\Lambda$ and $V_o$, and the NJL model plays therefore against an accelerating phase around the minimum of the potential, since $V_o$ is negative.

## 7. Dark Energy from NJL and SUSY Gauge Theory

As we have seen until now, the original NJL model has interesting cosmological consequences. However, the model by itself does not reproduce the observed feature of an accelerated expansion of the universe, and it is not desirable to introduce a cosmological constant by hand, without a good explanation. We rather ask for any model to be motivated from a deeper fundamental theory. Nowadays, a paradigm for such fundamental theory is played by Super Symmetric Field Theories, and a lot of work has been done in attempting to explain Dark Matter as well as Dark Energy as some super symmetric particle (references are given in the introduction, sec. 1). Nevertheless, any conclusive theory has been established yet to present date. We would like now to generalize the NJL potential to include a physically motivated potential from supersymmetric gauge theories. These class of models have been previously studied in Dark Energy models derived from gauge theory [28] [29] (and references therein), [21] [28]-[35] and are based on ADS (Affleck-Dine-Seiberg) superpotential [35]-[38]. The derived potential is of the form

$$U = \tilde{\Lambda}^{4+n}\phi^{-n}, \tag{48}$$

which is obtained from a non-perturbative super potential in a gauge theory, e.g. $n = 2\left(N_c + N_f\right)/\left(N_c - N_f\right)$ for an $SU\left(N_c\right)$ with $N_f$ flavours, and $\phi$ represent a fermion condensate, *i.e.* $\phi \approx \langle\bar{\psi}\psi\rangle$. The condensation energy $\tilde{\Lambda}$ is the scale of breaking of the gauge symmetry.[8] We now add the potential in Equation (48) to our NJL model. Since the effective NJL potential in Equation (13) has a quadratic term $\frac{1}{2}m^2\phi^2$ let us take $n = 2$ in Equation (48) so that we have the symmetry under $\phi \to 1/\phi$. Then, at some lower scale $\Lambda$, the self interaction of the field $\phi$ becomes more involved and the dynamics of the field is also governed by the effective NJL potential. By adding Equations (13) and (48), we would have the total potential

$$V = V_{NJL} + U = \frac{1}{2}m^2\phi^2 - \frac{\Lambda^4}{16\pi^2} f\left(x(\phi)\right) + \frac{\tilde{\Lambda}^6}{\phi^2}. \tag{49}$$

shown in Figure 12. Of course, this is an effective theory which is plausible to the ex-

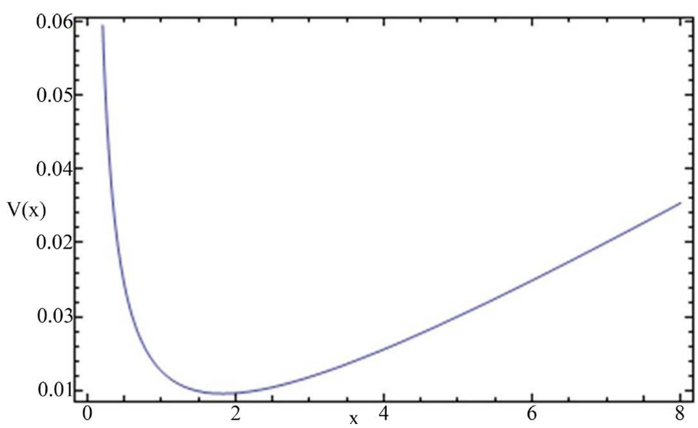



Figure 12. Graph of the total potential $V = V_{NJL} + U$, as a function of the variable $x$.

[8]In general, the scales $\tilde{\Lambda}$ and $\Lambda$ are not the same, and should not be confused.

tent that NJL and Equation (48) are valid or useful theories. This is on the same footing than using the NJL model to study the dynamics of hadrons, without having obtained the model directly from the QCD Lagrangian. Since we are simply adding a term to the already studied NJL potential, we use the results of the previous section 2. Using eq. ii), (17), the condition to be satisfied by the minimum $x$ is now written (remember that we wrote before $g = sg_c$, Equation (20))

$$1-\frac{1}{s^2}=-2x^2\left(\frac{4\pi^2 ms\tilde{\Lambda}^3}{\Lambda^4}\right)^2 - x\log\left(\frac{x}{1+x}\right)\equiv\alpha(x), \tag{50}$$

where we have defined the function $\alpha(x)$, which is seen to be parameterized through $\tilde{\Lambda},\Lambda,m$. This function has $-\infty<\alpha<+1$ and it is a monotonous growing function, regardless of the values of the parameters (they all are positive definite). Now, we know that the NJL potential $V_{NJL}$ is minimized in a non-trivial minimum $\phi_{\min}\neq 0$ when $g>g_c$, or $s>1$. In this case, the LHS in Equation (50) is a positive quantity, corresponding to $0<\alpha<1$ in the RHS, determining a solution $x=x_0$. This means that the total potential (49) still is minimized for some $x_0$, giving in its turn a non-trivial $\phi_0$. Given that the minimum $x_0$ satisfies (50), the minimized potential can be written

$$V_{\min}=\frac{\Lambda^4}{16\pi^2}\left[\frac{x_0}{s^2}-\log(1+x_0)\right]+\frac{6\pi^2}{x_0}\left(\frac{ms\tilde{\Lambda}^3}{\Lambda^2}\right)^2. \tag{51}$$

From this equation we see that, it is possible to obtain $V_{\min}>0$ (which would behave like a cosmological constant), if the parameters satisfy

$$\left(\frac{m\tilde{\Lambda}^3}{\Lambda^4}\right)^2>\frac{1}{6\left(4\pi^2\right)^2}\frac{1}{s^2}\left[x_0\log(1+x_0)-\frac{x_0^2}{s^2}\right]. \tag{52}$$

Let us show an example. Suppose that $g=\sqrt{2}g_c$, *i.e.* $s=\sqrt{2}>1$. Also, we need to say something about the parameters, so let us take $\Lambda^4=8\pi^2 m\tilde{\Lambda}^3$. In this way the Equation (50) is written

$$\frac{1}{2}=-\frac{1}{x^2}-x\log\left(\frac{x}{1+x}\right), \tag{53}$$

which has the solution $x=x_0\simeq 1.83$. Then, RHS Equation (52) gives the number $x_0\log(1+x_0)-\frac{x_0^2}{2}\simeq 0.23$. This means that, in order for the potential to be positive at the minimum, the parameters must satisfy $m\tilde{\Lambda}^3/\Lambda^4>1.22\times 10^{-5}$. We can use Equation (51) to obtain $V_{\min}$; in order of magnitude we have $V_{\min}\sim 10^{-2}\Lambda^4$. Let us now estimate some real physical values. The total energy density today is about $\rho_o=E_o^4\sim\left(10^{-3}\text{ eV}\right)^4$, and the Dark Energy contribution is $\rho_{DEo}=\Omega_{DEo}\rho_o$. If we identify our NJL fluid with DE, we would have $E_k+V(\phi)=\rho_{DEo}$. Now, in the limit of stabilized fields about the minimum[9] the energy density of our NJL fluid is $E_k+V(\phi)\to V_{\min}$. Then, (approximating $\Omega_{DEo}\simeq 2/3\sim 1$) we may write $\rho_{DEo}=E_0^4=V_{\min}$. Thus we may write

[9]We must keep in mind that in general, the field could be in a rolling regime, so the kinetic energy would not be negligible. Therefore we must be careful in the conditions that we are talking about.

$$\epsilon\Lambda^4 = 10^{-3}\ \text{eV} \tag{54}$$

with the precise value of coefficient $\epsilon$ depending on the value $V_{\min}$, as shown above (for our example we have $\epsilon \sim 10^{-2}$). Given that this theory allows $V_{\min} > 0$ as a result of the dynamics of the field, we have a possible explanation for the presence of a cosmological constant, and an accelerating universe.

## 8. Summary of Results and Discussion

The fermion model of Nambu and Jona-Lasinio (NJL) includes two different fermion states resulting from quantum effects, each one being associated with two different physical phases. For a weak coupling $g < g_c$ we have *massless fermion* fluid, whereas for a strong coupling $g > g_c$ a *massive fermion condensate* fluid is obtained. In this later case we can determine the mass of fermions and it is due to non-perturbative effects due to the strong coupling. A very convenient way to describe the system is to consider an equivalent scalar field $\phi$ moving under an effective potential $V = V_0^\phi + V_1^\phi$, which has a different form depending on the coupling strength.

Notice that in the strong coupling case $g > g_c$, the potential has a non-trivial negative minimum due to the negative contribution one-loop potential $V_1$ in Equation (6). The negative sign of this potential is due to the fermionic origin of $\psi$ field, and we have chosen to parameterize the fermion condensate in terms of an effective scalar field $\phi \sim \langle \bar{\psi}\psi \rangle$, as in Equation (4).

Here we studied the potential and solved the cosmological evolution for each fluid in presence of additional barotropic fluids (e.g. matter-dust or radiation).

For a weak coupling, we found a coefficient of state $\omega_\phi$ with oscillating values around zero, in such a way that the average value $\langle \omega_\phi \rangle = 0$. Also, because the potential goes as $V \sim \phi^2$ near the minimum, we have that the NJL fluid in the form of free fermions dilutes as a matter. A universe containing such a fluid (with or without matter and/or radiation) will expand forever without accelerating. On the other hand, a universe containing this NJL fluid besides a cosmological constant (with or without matter and/or radiation), will eventually accelerate necessarily, expanding forever.

On the other hand, the strong coupling case (without a cosmological constant) always causes an eventually vanishing energy density. This is due to the fact that the potential is negative when minimized, and even the additional presence of matter and/or radiation does not prevent this to happen. Since the vanishing energy (which is associated with the scale factor reaching a maximum), is followed by a contracting period, this means that a fermion condensate always makes the universe collapsed. The energy density of the field $\rho_\phi$ vanishes a couple of times (one in the expanding phase, and the another one in the contracting phase). Because of this, some quantities ($\Omega_\phi$, $\omega_\phi$) become inadequate to describe the fluid. It is important to point out the following interesting fact:

Equation (23) has been known and well studied since long time ago. If the curvature parameter is $k = +1$, the universe is said to have a *spherical* geometry; the scale factor is expected to get a null value eventually, so we have a collapsing universe. Because a spherical universe is also finite or *closed*, a collapsing universe was always associated with a closed universe. On the other hand, if $k = 0$, the universe has a *flat* geometry. For ordinary matter the total energy density could be diminishing, but it could never vanish effectively in a finite time, so the scale factor in this case is expected to be always increasing. Because divergent geodesic lines in a plane never meet again, a flat universe is said to be open. So, an open universe was thought to be infinite in size (although not necessarily, but in any case, always *growing*). Now, remember that from the beginning, in our present study, we have taken the curvature parameter to be $k = 0$, so we have been treating with a flat universe all the time. Nevertheless, we found that, if the universe contains a scalar field with a negative potential, then a future collapse cannot be avoided, giving a *collapsing flat* universe! In particular, because a negative potential arises naturally for the NJL model, a collapsing flat universe is also a *natural consequence*.

We also studied a variant of the strong coupling model, consisting in the addition of a cosmological constant. We found that, if the energy density $\rho_\Lambda$ is not big enough to overtake at least the minimized potential $V_{\min}$, the eventual receding of the scale factor cannot be avoided, and the universe will collapse inevitably. But if $\rho_\Lambda$ exceeds $V_{\min}$, then the scale factor will accelerate eventually, and the collapse will be absent.

Perhaps it is worth to emphasize that, in both cases of weak and strong coupling and *without* considering a cosmological constant, one may induce an acceleration of the scale factor by manipulating the initial condition for the field amplitude $\phi_i$, but we do not interest in it because 1) it has to be fine-tuned, and 2) it does not allow to include realistic models in which a previous deceleration period of matter dominance took place.

It is important to keep in mind that, once we settle a coupling strength (weak or strong), there is nothing in the theory to allow to switch between them, so actually a phase *transition* cannot be considered.

A very appealing feature of the NJL model is, in our opinion, the fact that 1) it is based on a "fundamental" symmetry (chiral symmetry), 2) the model leads to a potential which, due to quantum corrections, can adopt negative values in a natural way, and 3) it includes only one parameter: the coupling constant *g* (two parameters if we count the cut-off $\Lambda$). In return we obtain interesting consequences, as allowing more than one physical phase (each having different cosmological implications), and the possibility of a collapsing universe. This is to be compared with other models involving a symmetry breaking[10] or introducing new kinds of fluids aimed to be relevant to cosmological problems, but at the expense of introducing several fields or parameters.[11]

Finally, we saw that by considering an additional term besides the NJL potential, in

[10]For instance in Higgs-like models are required two parameters "$m$" and "$\lambda$" in order to get a potential $V = \frac{1}{2}m^2\varphi^2 + \frac{1}{4}\lambda\varphi^4$, which have to have a "correct" relation between them in order to break the symmetry.

[11]For instance, to "justify" the existence of scalar fields with useful potentials, frequently one has to invoke more sophisticated theories, like String, Kaluza-Klein, GUT's, etc. which demand a bigger effort to derive relevant results, and often implicate new exotic physics.

the form of an inverse power (which is motivated from some supersymmetric theories), then it is possible to obtain a total potential with a positive minimum, thus allowing to explain a cosmological constant as a consequence of a field dynamics, which is a fermion particle (instead of a scalar field) governed by simple basic symmetries.

## Acknowledgements

A.M. acknowledges financial support from UNAM PAPIIT Project No. IN101415 and Conacyt Fronteras Project No. 281.

## References


[1] Perlmutter, S., *et al.* (1999) *Astrophysical Journal*, **517**, 565-586. http://dx.doi.org/10.1086/307221

[2] Riess, A.G., *et al.* (1998) *Astronomical Journal*, **116**, 1009-1038. http://dx.doi.org/10.1086/300499

[3] Amanullah, R., *et al.* (2010) *Astrophysical Journal*, **716**, 712-738. http://dx.doi.org/10.1088/0004-637X/716/1/712

[4] Carlos, C., *et al.* (2010) *Astronomical Journal*, **139**, 519-539. http://dx.doi.org/10.1088/0004-6256/139/2/519

[5] Hicken, M., *et al.* (2009) *Astrophysical Journal*, **700**, 331-357. http://dx.doi.org/10.1088/0004-637X/700/1/331

[6] Riess, A.G., *et al.* (2007) *Astrophysical Journal*, **659**, 98-121. http://dx.doi.org/10.1086/510378

[7] Sakharov, A.D. (1966) *Soviet Physics—JETP*, **22**, 241.

[8] Peebles, P.J.E. and Yu, J.T. (1970) *Astrophysical Journal*, **162**, 815-836. http://dx.doi.org/10.1086/150713

[9] Sunyaev, R.A., *et al.* (1970) *Astrophysics and Space Science*, **7**, 3-19.

[10] Blake, C. and Glazebrook, K. (2003) *Astrophysical Journal*, **594**, 665-673. http://dx.doi.org/10.1086/376983

[11] Bennett, C.L., *et al.* (2013) *Astrophysical Journal Supplement*, **208**, 20. http://dx.doi.org/10.1088/0067-0049/208/2/20

[12] Ade, P.A.R., *et al.* (2015) arXiv:1502.01589.

[13] Reid, B.A., Chengalur, J.N., Begum, A. and Karachentsev, I.D. (2010) *Monthly Notices of the Royal Astronomical Society*, **404**, L60-L63. http://dx.doi.org/10.1111/j.1745-3933.2010.00835.x

[14] Percival, W.J., *et al.* (2001) *Monthly Notices of the Royal Astronomical Society*, **327**, 1297-1306. http://dx.doi.org/10.1046/j.1365-8711.2001.04827.x

[15] Tegmark, M., *et al.* (2001) *Monthly Notices of the Royal Astronomical Society*, **328**, 1039-1063. http://dx.doi.org/10.1046/j.1365-8711.2001.04902.x

[16] Padmanabhan, N., *et al.* (2012) *Monthly Notices of the Royal Astronomical Society*, **427**, 2132-2145. http://dx.doi.org/10.1111/j.1365-2966.2012.21888.x

[17] Anderson, L., *et al.* (2014) *Monthly Notices of the Royal Astronomical Society*, **441**, 24-62. http://dx.doi.org/10.1093/mnras/stu523

[18] Eisenstein, D.J., *et al.* (2005) *Astrophysical Journal*, **633**, 560-574. http://dx.doi.org/10.1086/466512

[19] Tojeiro, R., *et al.* (2014) *Monthly Notices of the Royal Astronomical Society*, **440**, 2222-2237. http://dx.doi.org/10.1093/mnras/stu371

[20] Copeland, E.J., Sami, M. and Tsujikawa, S. (2006) *International Journal of Modern Physics D*, **15**, 1753-1935. http://dx.doi.org/10.1142/S021827180600942X

[21] Ratra, B. and Peebles, P.J.E. (1988) *Physical Review D*, **37**, 3406. http://dx.doi.org/10.1103/PhysRevD.37.3406

[22] Steinhardt, P.J., Wang, L. and Zlatev, I. (1999) *Physical Review Letters*, **82**, 896. http://dx.doi.org/10.1103/PhysRevLett.82.896

[23] Steinhardt, P.J., Wang, L. and Zlatev, I. (1999) *Physical Review D*, **59**, Article ID: 123504. http://dx.doi.org/10.1103/PhysRevD.59.123504

[24] De la Macorra, A. and Piccinelli, G. (2000) *Physical Review D*, **61**, Article ID: 123503.

[25] De la Macorra, A. and Stephan-Otto, C. (2002) *Physical Review D*, **65**, Article ID: 083520. http://dx.doi.org/10.1103/PhysRevD.65.083520

[26] Das, S., Corasaniti, P.S. and Khoury, J. (2006) *Physical Review D*, **73**, Article ID: 083509.

[27] De la Macorra, A. (2007) *Physical Review D*, **76**, Article ID: 027301. http://dx.doi.org/10.1103/PhysRevD.76.027301

[28] De la Macorra, A. (2008) *Journal of Cosmology and Astroparticle Physics*, **0801**, 030.

[29] De la Macorra, A. (2007) *Astroparticle Physics*, **28**, 196-204. http://dx.doi.org/10.1016/j.astropartphys.2007.05.005

[30] De la Macorra, A. (2004) *Physics Letters B*, **585**, 17-23. http://dx.doi.org/10.1016/j.physletb.2004.02.006

[31] De la Macorra, A. (2010) *Astroparticle Physics*, **33**, 195-200. http://dx.doi.org/10.1016/j.astropartphys.2010.01.009

[32] De la Macorra, A. (2005) *Physical Review D*, **72**, Article ID: 043508.

[33] De la Macorra, A. (2003) *Journal of High Energy Physics*, **0301**, 033.

[34] De la Macorra, A. and Stephan-Otto, C. (2001) *Physical Review Letters*, **87**, Article ID: 271301. http://dx.doi.org/10.1103/PhysRevLett.87.271301

[35] Binetruy, P. (1999) *Physical Review D*, **60**, Article ID: 063502. http://dx.doi.org/10.1103/PhysRevD.60.063502

[36] Masiero, A., Pietroni, M. and Rosati, F. (2000) *Physical Review D*, **61**, Article ID: 023509.

[37] Lyth, D.H. and Riotto, A. (1999) *Physics Reports*, **314**, 1-146. http://dx.doi.org/10.1016/S0370-1573(98)00128-8

[38] Affleck, I., Dine, M. and Seiberg, N. (1985) *Nuclear Physics B*, **256**, 557-599. http://dx.doi.org/10.1016/0550-3213(85)90408-0

Journal of Modern Physics, 2016, 7, 1801-1810
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Fundamental Harmonic Power Laws Relating the Frequency Equivalents of the Electron, Bohr Radius, Rydberg Constant with the Fine Structure, Planck's Constant, 2 and π

**Donald William Chakeres**

Department of Radiology, The Ohio State University, Columbus, OH, USA
Email: donald.chakeres@osumc.edu

**How to cite this paper:** Chakeres, D.W. (2016) Fundamental Harmonic Power Laws Relating the Frequency Equivalents of the Electron, Bohr Radius, Rydberg Constant with the Fine Structure, Planck's Constant, 2 and π. *Journal of Modern Physics*, **7**, 1801-1810.
http://dx.doi.org/10.4236/jmp.2016.713160

**Received:** August 3, 2016
**Accepted:** September 27, 2016
**Published:** September 30, 2016



## Abstract

We evaluate three of the quantum constants of hydrogen, the electron, $e^-$, the Bohr radius, $a_0$, and the Rydberg constants, $R_\infty$, as natural unit frequency equivalents, $v$. This is equivalent to Planck's constant, $h$, the speed of light, $c$, and the electron charge, $e$, all scaled to 1 similar in concept to the Hartree atomic, and Planck units. These frequency ratios are analyzed as fundamental coupling constants. We recognize that the ratio of the product of $8\pi^2$, the $v_{e^-}$ times the $v_R$ divided by $v_{a0}$ squared equals 1. This is a power law defining Planck's constant in a dimensionless domain as 1. We also find that all of the possible dimensionless and dimensioned ratios correspond to other constants or classic relationships, and are systematically inter-related by multiple power laws to the fine structure constant, $\alpha$; and the geometric factors 2, and π. One is related to an angular momentum scaled by Planck's constant, and another is the kinetic energy law. There are harmonic sinusoidal relationships based on 2π circle geometry. In the dimensionless domain, $\alpha$ is equivalent to the free space constant of permeability, and its reciprocal to permittivity. If any two quanta are known, all of the others can be derived within power laws. This demonstrates that $8\pi^2$ represents the logical geometric conversion factor that links the Euclid geometric factors/three dimensional space, and the quantum domain. We conclude that the relative scale and organization of many of the fundamental constants even beyond hydrogen are related to a unified power law system defined by only three physical quanta of $v_{e^-}$, $v_R$, and $v_{a_0}$.

## Keywords

Fundamental Physical Constants, Unification Models, Hydrogen, Electron, Bohr Radius, Rydberg Constant, Fine Structure Constant

## 1. Introduction

The quantum properties characterizing hydrogen are fundamental constants important in many divergent areas of physics [1]. We focus on the electron, $e^-$, Bohr radius, $a_0$, and the Rydberg constant, $R_\infty$, within hydrogen as a unified physical and mathematical system. These constants represent a mass, a distance, and *a* frequency (1/time) so they span the physical units. These constants are all known to be inter-related through power laws with the geometric factors of 2, π, the fine structure constant, $\alpha$, the speed of light, $c$, and Planck's constant, $h$. In fact, each one of the hydrogen quantum constants is defined utilizing $\alpha$, $c$, $h$ in its derived unit value. None of them are experimentally directly measurable. This system is associated with all of the elements through Mosley's law. These quanta are related to the kinetic energy equation through Sommerfeld's original physical interpretation of $\alpha$ as a velocity divided by $c$, a dimensionless $\beta$. For the hydrogen system, the product of the rest mass of the electron times the speed of light times $\alpha$ both squared divided by 2 equals the hydrogen ionization energy. Through $\alpha$ and these constants, broad segments of physics are directly mathematically and conceptually inter-related within systematic power laws [2] [3].

Our goal is to evaluate these three quantum values as frequency equivalents: $\nu_R$, $\nu_{a_0}$, and $\nu_{e^-}$ in the simplest fundamental manner, ratios, which are independent of any arbitrary physical unit system. We evaluate these three quantum constants all as natural unit frequency equivalents, $\nu$ Hz similar to Planck's time, and the Hartree atomic units [4]-[6]. This unit transformation is equivalent to Planck's constant, $h$, the speed of light, $c$, the electron charge, $e$, all scaled to 1. By definition when $h$ equals 1, its angular momentum is dimensionless. Their SI unit values can be reconstructed by multiplying the dimensionless calculation times the SI units. The final results are identical. We are searching for power law and harmonic sinusoidal patterns since these quanta are associated with quantum oscillators that have wave properties, and these constants are inter-related by many different power laws utilizing their SI units. Any other physical units such as mass, energy, or distance would result in the identical coupling constants, and not change the conclusions or dimensionless calculations. We conceptually choose frequency as our physical unit of choice since we are searching for wave properties.

We find that these three quantum ratios and $\alpha$ represents an integrated power law system broader than presently recognized. They are fundamentally related to the geometric factors of 2 and π as a composite in $8\pi^2$ [7]. It is possible to define any of these quantum values as a power law of any other quantum value utilizing these geometric factors. We also find that many other fundamental constants can be simply defined in terms of these three quanta in this dimensionless domain simplifying and unifying the unit system greatly.

## 2. Methods

All data for the constants' transformations to frequency equivalents were obtained from the websites: http://physics.nist.gov/cuu/Constants/, and
http://physics.nist.gov/cuu/Constants/energy.html The NIST site has an online physical

unit converter that can be used for these types of calculations. The respective frequency equivalents are: the Rydberg constant, $\nu_R$, 3.28984196(17) × $10^{15}$ *Hz*, Bohr radius, $\nu_{a_0}$, 5.66525639(28) × $10^{18}$ *Hz*, and the electron, $\nu_{e^-}$, 1.23558996(05) × $10^{20}$ *Hz*. The reciprocal of $\alpha$ is 137.0359999(78). The relative precision of the calculations is 5 × $10^{-8}$, and limited by the uncertainty of *h*.

The italic number 1 is utilized after the standard constant's symbol to notate that is it scaled in the units of 1 for *h*, *c*, and *e* rather than SI units, or in the dimensionless domain. Two examples are $\varepsilon_0 1$, and *h*1. Many of the possible dimensionless and dimensional ratios of $\nu_R$, $\nu_{a_0}$, and $\nu_{e^-}$ define fundamental integer power laws, and harmonic circle sinusoidal relationships.

We use a specialized notation to abbreviate the power law ratio relationships. These all represent fundamental constants. The notation allows for complicated ratios and powers to be expressed as text. Fundamental constants are commonly associated with super and subscript notations characterizing them. Capital *A* is chosen to represent ratios. The superscript are the symbols for the frequencies of $\nu_R$, $\nu_{a_0}$, $\nu_{e^-}$ or $\alpha$ of the numerator. The subscript is for the denominator. The powers of $\nu_R$, $\nu_{a_0}$, $\nu_{e^-}$ or $\alpha$ are within parentheses to their right. These represent power laws so different points along a single power line are all inter-related. For example, $A_{\nu_{e^-}(14)}^{\nu_{a_0}(15)}$ *Hz*, 1.028140195(51) *Hz*, and $A_{\nu_{a_0}(15)}^{\nu_{e^-}(14)}$ *s*, 0.972630001(49) *s* are reciprocals. All three of the following *A* values fall on a common power law line, $A_{\nu_{e^-}(14)}^{\nu_{a_0}(15)} Hz$, $A_{\nu_{e^-}}^{\nu_{a_0}(15/14)} Hz^{(1/14)}$ 1.00198421(05) $Hz^{(1/14)}$, and $A_{\nu_{e^-}(14/15)}^{\nu_{a_0}} Hz^{(1/15)}$, 1.00185181(05) $Hz^{(1/15)}$. Note that within the dimensional domain the *A* powers related to *Hz* vary, but add to an integer power of *Hz* when transformed back to the SI units.

Equation (1a) demonstrates a pure geometric factor. $1/8\pi^2$ equals the ratio of the product of $\nu_{e^-}$ and $\nu_R$ divided by the square of $\nu_{a_0}$ [7]. This is the foundation of all of the other power laws. Different arrangements of equivalent relationships are shown as lettered Equations. Equation (1b) rearranges the equation to an identity Equation equaling 1.

$$\frac{1}{8\pi^2} = \left(\frac{\nu_{e^-}\nu_R}{\nu_{a_0}^2}\right) = A_{\nu_{a_0}(2)}^{\nu_{e^-}\nu_R} = A_{\nu_{a_0}}^{\nu_{e^-}} A_{\nu_{a_0}}^{\nu_R} \tag{1a}$$

$$1 = 8\pi^2\left(\frac{\nu_{e^-}\nu_R}{\nu_{a_0}^2}\right) = 8\pi^2 \frac{1}{2\pi\alpha}\frac{\alpha}{4\pi} = 8\pi^2 A_{\nu_{a_0}(2)}^{\nu_{e^-}\nu_R} = 8\pi^2 A_{\nu_{a_0}}^{\nu_R} A_{\nu_{a_0}}^{\nu_{e^-}} \tag{1b}$$

Equations (2a), (2b), (3a, 3b), (4a), (4b), (4c) are all of the possible dimensionless frequency ratios of $\nu_R$ $\nu_{a_0}$, and $\nu_{e^-}$; and are systematically related to $\alpha$ and the geometric factors: 2, 2π, and 4π [7]. Each represents another fundamental constant or relationship. $A_{\nu_{e^-}}^{\nu_{a_0}} 1$ is related to the ratio of $\nu_{a_0}$ divided by $\nu_{e^-}$, or $2\pi\alpha$, and equals 0. 0458506184(23). $A_{a_0}^{R} 1$ equals $\nu_R$ divided by $\nu_{a_0}$, $\alpha/4\pi$, and equals 5.80704866(29) × $10^{-4}$. $A_{\nu_{e^-}}^{R} 1$ equals $\nu_R$ divided by $\nu_{e^-}$, $\alpha^2/2$, and equals 2.66256772(13) × $10^{-5}$.

$$\left(\frac{\nu_{a_0}}{2\pi\nu_{e^-}}\right) = \alpha \tag{2a}$$

$$\left(\frac{\nu_{a_0}}{\nu_{e^-}}\right) = 2\pi\alpha = A_{\nu_{e^-}}^{\nu_{a_0}} 1 \tag{2b}$$

$$\left(\frac{4\pi\nu_R}{\nu_{a_0}}\right) = \alpha \tag{3a}$$

$$\left(\frac{\nu_R}{\nu_{a_0}}\right) = \frac{\alpha}{4\pi} = A_{\nu_{a_0}}^{\nu_R} 1 \tag{3b}$$

$$\left(2\nu_R/\nu_{e^-}\right)^{1/2} = \alpha \tag{4a}$$

$$\frac{\nu_R}{\nu_{e^-}} = \frac{\alpha^2}{2} = A_{\nu_{e^-}}^{\nu_R} 1 \tag{4b}$$

$$\nu_R = \frac{\alpha^2}{2}\nu_{e^-} = \nu_{e^-} A_{\nu_{e^-}}^{\nu_R} 1 \tag{4c}$$

The derivation of the dimensionless values for permittivity, $\varepsilon_0 1$ and permeability, $\mu_0 1$, are shown in Equations (5)-(11). Here $\varepsilon_0 1$ equals $1/\alpha$, 137.0359999(78), and permeability, $\mu_0 1$ equals $\alpha$, 7.29735256(36) × $10^{-3}$. Equation (5) demonstrates the relationship between $\alpha$, electron charge, $e$, $\varepsilon_0$, 2, 4π, $h$, $\hbar$, and $c$ used in the derivation. Equation (6) derives the ionization energy of the electron in hydrogen. Equation (7) transforms the ionization energy from joules to frequency in $Hz$ by dividing by $h$. In Equation (7) the $a_0$ is converted to a frequency equivalent utilizing a ratio with $c$. Equation (8) transforms the relationship to the dimensionless coupling constant of $\nu_R$ divided by $\nu_{a_0}$. Substitution of Equation (5) converts to the ratio of $\alpha$ and 4π, identical to Equation (3b).

$$\alpha = \left(\frac{e^2}{2\varepsilon_0 hc}\right) = \frac{1}{4\pi\varepsilon_0}\left(\frac{e^2}{\hbar c}\right) = \left(\frac{k_e e^2}{\hbar c}\right) \tag{5}$$

$$E_R = \left(\frac{1}{4\pi\varepsilon_0}\right)\left(\frac{e^2}{2a_0}\right) \tag{6}$$

$$\nu_R = \frac{E_R}{h} = \left(\frac{e^2\nu_{a_0}}{8\pi\varepsilon_0 hc}\right) \tag{7}$$

$$\frac{\nu_R}{\nu_{a_0}} = \left(\frac{e^2}{8\pi\varepsilon_0 hc}\right) = \frac{1}{4\pi}\left(\frac{e^2}{2\varepsilon_0 hc}\right) = \left(\frac{\alpha}{4\pi}\right) = A_{\nu_{a_0}}^{\nu_R} 1 = k_e 1 \tag{8}$$

Coulomb's dimensionless unit 1 constant, $k_e 1$, must equal $\alpha/4\pi$, 5.80704866(29) × $10^{-4}$. Equations (8) and (9) derive the dimensionless $\varepsilon_0 1$ constant from $k_e 1$. Here, $\varepsilon_0 1$ must equal $1/\alpha$, 137.0359999(78).

$$k_e 1 = \left(\frac{1}{4\pi(\varepsilon_0 1)}\right) = \left(\frac{1}{4\pi\left(\frac{1}{\alpha}\right)}\right) = \left(\frac{\alpha}{4\pi}\right) = \frac{\nu_R}{\nu_{a_0}} = A_{\nu_{a_0}}^{\nu_R} 1 \tag{9}$$

The dimensionless relationship of $\varepsilon_0 1$, $\mu_0 1$, and $c1$ is shown in Equation (10). $\mu_0 1$ equals $\alpha$, 7.29735256(36) × 10^{−3}.

$$\mu_0 1 \varepsilon_0 1 c1^2 = \alpha\left(\frac{1}{\alpha}\right)c1^2 = 1 \tag{10}$$

The following is an example a power law derivation of one hydrogen quantum value from another utilizing a different, but parallel method. This example derives $\nu_{a_0}$ from $\nu_R$, $\alpha$, and 4π based on Equation (3). Equation (11) evaluates the ratio of $4\pi/\alpha$ or $A_{\nu_R}^{\nu_{a_0}}$ raised to a consecutive integer power series divided by $\nu_R$ in search of the value in seconds that is closest to 1 second. All of the constants demonstrate this power pattern. This is equivalent to when $4\pi/\alpha$ or $A_{\nu_R}^{\nu_{a_0}}$ raised to a power is nearly equal to $\nu_R 1$. This can be substituted for a $\nu 1$ value. The powers of $\nu_{a_0}$ and $\nu_R$ must be separated by 1, so the ratio is not dimensionless. The $A$ values are related to identity relationship within these equations, but these $A$ values represent fundamental constants. These powers represent natural quantum number powers. They can be derived by searching arbitrarily through ratio power matrix of consecutive integer series for each constant's power. The ratio where the scalar is closest to 1 represents the $A$ ratio and the natural powers. Table 1 is an example of the type of arbitrary search for natural powers linking $\nu_R$ and $\nu_{a_0}$ All of the constants demonstrate this pattern with scalars near 1.

$$\frac{\left(\frac{4\pi}{\alpha}\right)^n}{\nu_R}s = \frac{\left(\frac{\nu_{a_0}}{\nu_R}\right)^n}{\nu_R}s = \left(\frac{\nu_{a_0}{}^n}{\nu_R{}^{n+1}}\right)s = A_{\nu_R(n+1)}^{\nu_{a_0}(n)} \approx 1s \tag{11}$$

For these two hydrogen constants, the $A_{\nu_R(n+1)}^{\nu_{a_0}(n)}$ value closest to 1s occurs when the $n$ power is 5. The ratio of $4\pi/\alpha$ to the fifth power, equals $\nu_R$ times $\nu_{a_0}^5$ divided by $\nu_R^6$, or the product of $\nu_R$ and $A_{\nu_R(6)}^{\nu_{a_0}(5)}s$, Equation (12). Here, $4\pi/\alpha$ to the fifth power equal 1.514337482(75) × 10^{16}.

$$\left(\frac{4\pi}{\alpha}\right)^5 = \nu_R Hz\left(\frac{\nu_{a_0}^5}{\nu_R^6}\right)s = \nu_R Hz A_{\nu_R(6)}^{\nu_{a_0}(5)}s = \nu_R Hz \times 4.60307060(23)s \tag{12}$$

Table 1. Power law ratio matrix.

| $\nu_R^{row}/\nu_{a_0}^{column}$ | 4 | 5 | 6 | 7 |
|---|---|---|---|---|
| 4 | $1.13 \times 10^{-13}$ | $2.00 \times 10^{-32} s$ | $3.54 \times 10^{-51} s^2$ | $6.25 \times 10^{-70} s^3$ |
| 5 | $3.74 \times 10^{2}\ Hz$ | $6.60 \times 10^{-17}$ | $1.16 \times 10^{-35} s$ | $2.05 \times 10^{-54} s^2$ |
| 6 | $1.23 \times 10^{18}\ Hz^2$ | $2.17 \times 10^{-1}\ Hz$ | $3.83 \times 10^{-20}$ | $6.76 \times 10^{-39} s$ |
| 7 | $4.04 \times 10^{33}\ Hz^3$ | $7.14 \times 10^{14}\ Hz^{14}$ | $1.26 \times 10^{-4}\ Hz$ | $2.22 \times 10^{-23}$ |

Table 1 is a power ratio matrix of $\nu_R$ raised to the row power divided by $\nu_{a0}$ raised to the column power where the powers are consecutive integer series. The ratio with a scalar value closest to 1 is searched for within an arbitrary matrix. Those powers represent the natural quantum number powers linking those two constants. This occurs with $\nu_R$ raised to the 6th power divided by $\nu_{a0}$ raised to the 5th power. The ratio is $2.17 \times 10^{-1}$ *Hz*. Note that all of the other values are widely divergent from 1. Here $2.17 \times 10^{-1}$ *Hz* is the reciprocal of $A_{\nu_R(6)}^{\nu_{a_0}(5)}s$, and the reciprocal of 4.06 *s* as seen in Equations (12)-(15).

The ratio relationships in Equations (2)-(4) are solved for any hydrogen quantum value, and are raised to the ninth power which is closest to a scalar value of 1, for this example $\nu_{a_0}$, and the power of 5. This exposes the powers of Hz and seconds, Equation (13). The derived value for $4\pi/\alpha$ raised to the fifth power is substituted from Equation (12) into Equation (13). The fifth power of $\nu_{a_0}$ is derived from the product of $\nu_R$ raised to the sixth power and $A_{\nu_R(6)}^{\nu_{a_0}(5)} s$.

$$\nu_{a_0}^5 = \left(\frac{4\pi}{\alpha}\right)^5 \nu_R^5 = \nu_R^6 Hz^6 \left(\frac{\nu_{a_0}^5}{\nu_R^5}\right) s = \nu_R^6 Hz^6 A_{\nu_R(6)}^{\nu_{a_0}(5)} s \tag{13}$$

In Equation (14) $\nu_{a_0}$ is derived from fractional powers of $\nu_R$. Here, $A_{\nu_R(6/5)}^{\nu_{a_0}} s^{(1/5)}$ equals 1.35709274(07) $s^{(1/5)}$.

$$\nu_{a_0} = \left(\frac{4\pi}{\alpha}\right)\nu_R = \nu_R^{6/5} Hz^{6/5} \left(\frac{\nu_{a_0}}{\nu_R^{6/5}}\right) s^{(1/5)} = \nu_R^{6/5} Hz^{6/5} A_{\nu_R(6/5)}^{\nu_{a_0}} s^{(1/5)} \tag{14}$$

Equation (15) is another power law variation of this relationship deriving $\nu_{a_0}$ from $\nu_R$. $A_{\nu_{a_0}(5)}^{\nu_R(6)} Hz$ equals 0.21724628739519 Hz. The sixth power of $4\pi/\alpha$ equals 2.60775752(14) × $10^{19}$.

$$\nu_{a_0} = \left(\frac{(4\pi)^6}{\alpha^6}\right)\left(\frac{(\nu_R)^6}{(\nu_{a_0})^5}\right) Hz = \left(\frac{(4\pi)^6}{\alpha^6}\right) A_{a_0(5)}^{R(6)} Hz \tag{15}$$

Equation (16) is another example that derives $\nu_{e^-}$ from $\nu_R$. Here, $A_{\nu_R(9/7)}^{\nu_{e^-}} s^{(2/7)}$ equals 1.38426983(05) $s^{(2/7)}$.

$$\nu_{e^-} = \nu_R^{(9/7)} Hz^{(9/7)} \left(\frac{\nu_{e^-}}{\nu_R^{(9/7)}}\right) s^{(2/7)} = \nu_R^{(9/7)} Hz^{(9/7)} A_{\nu_R(9/7)}^{\nu_{e^-}} s^{(2/7)} \tag{16}$$

## 3. Results

Equation (1) is related to Planck's constant in the dimensionless domain, and equals 1, Equations (17a), (17b). This is logical since $h1$ is intentionally scaled as 1 in this system since energy and frequency are scaled identically in a pure frequency domain. This ratio is equivalent to an angular momentum since the ratio represents the product of a mass, a frequency, a distance squared, (kg·m²/s) in the SI dimensional domain. Since this ratio is in the frequency domain, a distance is related to the reciprocal of the frequency equivalent. This relationship is related to the annihilation energy of a mass. Equation (17b) demonstrates that the reduced Planck's constant, $\hbar$, equals 1/2 in the dimensionless domain, $\hbar 1$.

$$h1 = 8\pi^2 \left(\frac{\nu_{e^-}\nu_R}{\nu_{a_0}^2}\right) = 8\pi^2 A_{\nu_{a_0}(2)}^{\nu_{e^-}\nu_R} = 8\pi^2 \frac{1}{2\pi\alpha}\frac{\alpha}{4\pi} = 8\pi^2 A_{\nu_{a_0}}^{\nu_R} A_{\nu_{a_0}}^{\nu_{e^-}} = 1 \tag{17a}$$

$$\hbar 1 = 4\pi\left(\frac{(\nu_R s)(\nu_{e^-} s)}{(\nu_{a_0})^2}\right) = \frac{1}{2} \tag{17b}$$

Equations (18a) and (18b) demonstrate that this same relationship is valid in the standard SI unit equations of $h$ and $\hbar$.

$$h = 8\pi^2 m_{e^-} v_R a_0^2 \tag{18a}$$

$$\hbar = 4\pi m_{e^-} v_R a_0^2 \tag{18b}$$

Equation (4c) is equivalent to the prototype kinetic energy equation where the product of a mass times the velocity divided by $c$, $\beta$ squared divided by 2.

Equation (2) demonstrates that the $\alpha$ derivation from $v_{a_0}$ and $v_{e^-}$ is within a circle equation related to a radius, circumference, and 2π. This must be related to a sinusoidal harmonic system, and a fundamental momentum relationship.

The dimensionless constants that define $\varepsilon_0 1$, $\mu_0 1$, and $k_e 1$ are all very simple values related to $\alpha$ and geometric factors, Equations (8)-(10). $k_e 1$ logically is related to $\alpha/(4\pi)$ since the energy is proportional to the distance frequency equivalent. This is a universal relationship for any distance as a frequency equivalent. Here, $v_{a_0}$ times $\alpha/(4\pi)$ equals $v_R$, Equation (3).

Equations (12)-(16) demonstrate that if any two of the quantum values, including an $A$ factor, are known then all of the constants can be derived.

The $A$ factors do not represent errors, but are fundamental essential constants incorporating the geometric factors that are imbedded within the quantum domain scalar values. These geometric factors represent conversion factors that bridge between Euclid geometry; and the power laws of $v_R$, $v_{a_0}$, $v_{e^-}$ and $\alpha$. The geometric factors are essential in projecting the power laws into the harmonic linear domain, and three dimensional spaces from a system solely defined by sinusoidal waves. Here, 2 is related to kinetic energy, Equation (4). In Equations (17a) and (17b), the factor 2 is similar to the 2 in the Schwarzschild equation since it refers to the transition of matter to an annihilation boson state. This is not a kinetic state. Here, 2π is related to a radius and its associated circumference as seen in Equations (17a) and (17b). Here, 4π is related to 4π times the radius squared that defines the surface of the sphere seen in force fields such as electrical charges, or magnetic fields, as in $k_e 1$, Equations (3), (5), (7)-(9), (11)-(15).

## 4. Discussion

The existing physical unit system of SI units is an arbitrary system of mass, time, distance, and energy units. They were chosen for measurement convenience, and their previous utilization/standardization. In quantum physics integer quantum numbers are associated with natural quantum units, such as the Rydberg constant. Planck's time is a classic time natural unit. Quantum spin is associated with $h$ as a natural unit. This method has a long history, and is a valid approach in physics [4]-[6]. Natural units are not arbitrary, and essential if the system is to be evaluated from a quantum number perspective. The ratio of two natural units of a single physical unit represents a fundamental dimensionless scaling ratio. The Standard Model is not helpful in most scaling relationships between particles. The Higgs boson could not be predicted with accuracy within the Standard Model. The Buckingham Pi theory of dimensionless analysis of physical systems is well established, and states that physical systems can be analyzed in a dimensionless domain. These reasons are the rationale for the methods utilized to search for the most fundamental quantum relationships.

This paper demonstrates that many of the fundamental constants in this hydrogen system demonstrate very simple, and integrate definitions as ratios of three hydrogen quanta and geometric factors. It is logical that all of these ratios should be related to fundamental constants and relationships. These are truly the only three constants in this global power law system. This is not the typical concept of these individual constants. Though they are inter-related they are not thought of as being only ratio or product projections of just three quanta. Equations (1), (17a), (17b) demonstrate that Planck's constant is related to the dimensionless ratio of the frequencies defining an angular momentum scaled to the number 1. This is anticipated in a natural unit system. Planck's reduced, $\hbar 1$, constant is scaled to 1/2. This is logical since Planck's constant times one Hz represents at minimum unit of energy per cycle. This equates the annihilation energy of a mass to an equivalent frequency. The $h1$ is directly related to a spin of 1 since it refers to 1 Hz times a dimensionless value associated with the energy of a wave with a frequency of 1 Hz. Photons have a spin of 1. The reduced Planck's constant $\hbar 1$ is 1/2, and associated with a spin of 1/2 of a fermion, in this case the electron. The $8\pi^2$ geometric factor is the composite of $2\pi$ squared, $4\pi^2$; and 2. The $2\pi$ squared scales the circumference distance an electron orbiting the proton. The 2 is related to the fact this angular momentum refers to the process of mass annihilation to photons, not to a kinetic energy. $8\pi^2$ is also found as the geometric factor of the Schrodinger equation.

Multiple different ratios of $\nu_R$, $\nu_{a_0}$, $\nu_{e^-}$, and geometric factors are related to $\alpha$, Equations (1)-(4). One of them is equivalent to the kinetic energy of a mass as a prototype form for all kinetic energies. One of the ratios is related to $2\pi$ a relationship of $\nu_{a_0}$ and $\nu_{e^-}$, Equation (2). This demonstrates a harmonic sinusoidal nature to this system. This must also define a fundamental relationship between distance and mass, (Equation (2)) similar to that seen with distance and energy in Equation (3). This must be related to a linear momentum. The Coulomb constant's dimensionless value, $k_e 1$ is related to the ratio of $\nu_R$ divided by $\nu_{a_0}$, or $\alpha/4\pi$. Here $4\pi$ is the classic geometric factor related to a spherical surface and a force field. Dimensionless permittivity and permeability equal $1/\alpha$, and $\alpha$. This is a remarkably unified and simplified system of units demonstrating their most fundamental scaling inter-relationships.

This work demonstrates that there are a myriad of power law relationships where any constant can derived if any two are known including an $A$ value, Equations (12)-(16). There are many power laws of these constants in the SI unit domain, but since their scaling relationships are not apparent these types of derivations starting with only three quanta are not possible. We present only a few of the possible power laws. In the SI unit system there are many different constants that appear to be unique values, but based on this analysis they actually represent only a product or ratio combination of the same three hydrogen quanta. Equation (5) is a standard unit equation and inter-relates five different physical constants. In this dimensionless unified domain three of the constants drop out from the scaling of the calculations as 1's. The other two are inter-related by a geometric factor. Two of the constants are both related to the same natural unit value, $\alpha$. There is really only one essential quantum constant that inter-relates all five of the SI values.

These power law relationships are not felt to coincidental or mathematically contrived, but rather fundamental. The ratio of the $\nu_{a_0}$ to the 15 power divided by $\nu_{e^-}$ to the 14 power equals 1.028140195(51) *Hz*. This is a ratio of $1.987438 \times 10^{281}$ $Hz^{15}$ divided by $1.933042 \times 10^{281}$ $Hz^{14}$. It is highly unlikely that ratios of the integer powers of these quantum values, that represent gigantic scalar values far beyond what is typical, are all nearly equal to 1 by chance. Every hydrogen quantum constant demonstrates a similar pattern. It is also essential that they not be equal to 1 so that is not a logic or mathematical error. They cannot be equal to 1 since the geometric factors are imbedded in their scalar values. There are many similar fundamental ratio power laws, and most are not known. A recent paper describes the quantum mechanical derivation of the Wallis formula for π [8]. Therefore, these types of relationships between the quantum constants and geometric factors of 2 and π do not represent inappropriate speculation, but are logically imbedded within the constants' natural unit scaling.

## 5. Conclusion

This paper describes the fundamental scaling relationships between the unified physical system of the electron, Bohr radius, and Rydberg constant of hydrogen within a natural unit dimensionless or dimensional system. It is found that classic geometric factors are embedded within the scaling of these quanta. This system bridges from the quantum wave domain to three dimensional space domains. When the relationships of these three constants are analyzed, their simple ratios and products project out to a wide array of other fundamental constants and relationships. There are also a myriad of power laws, and $A$ values, some known and some new, inter-relating these constants so that they can all be derived from the knowledge of just two. These finding demonstrate that the fundamental constants of hydrogen represent a highly integrated logical harmonic power law system that extends beyond hydrogen. This represents a new perspective on the Standard Model within a parallel power law system.

## Acknowledgements

I would like to thank Richard White M. D., Vola Andrianarijaon PhD, Keshab Sapkota PhD, and Richard Vento MS for their help.

## References

[1] Zhao P., Lichten, W. and Zhou, Z. (1989) *Physical Review A*, **29**, 2888. http://dx.doi.org/10.1103/PhysRevA.39.2888

[2] Saichev, A., Malevergne, Y. and Sornette, D. (2009) *Lecture Notes in Economics and Mathematical Systems*, **632**.

[3] Katznelson, Y. (2004) An Introduction to Harmonic Analysis. 3rd Edition, Cambridge

University Press, Cambridge. http://dx.doi.org/10.1017/CBO9781139165372

[4] Mills, J. (1932) *The Journal of Physical Chemistry*, **36**, 1089-1107. http://dx.doi.org/10.1021/j150334a001

[5] Karshenboima, S. (2008) *The European Physical Journal Special Topics*, **163**, 141.

[6] Flowers, J. and Petley, B. (2001) *Reports on Progress in Physics*, **64**, 1191. http://dx.doi.org/10.1088/0034-4885/64/10/201

[7] Chakeres, D. (2011) *Particle Physics Insights*, **4**, 33. http://dx.doi.org/10.4137/PPI.S8269

[8] Friedmann, T. and Hagen, C. (2015) *Journal of Mathematical Physics*, **56**, 112101. http://dx.doi.org/10.1063/1.4930800

Journal of Modern Physics, 2016, 7, 1811-1817
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# Oscillation of Davydov Solitons in Three Wells

**Bi Qiao**

Wuhan University of Technology, Wuhan, China
Email: biqiao@gmail.com

**How to cite this paper:** Qiao, B. (2016) Oscillation of Davydov Solitons in Three Wells. *Journal of Modern Physics*, **7**, 1811-1817.
http://dx.doi.org/10.4236/jmp.2016.713161

**Received:** August 3, 2016
**Accepted:** September 27, 2016
**Published:** September 30, 2016



## Abstract

In this work, we propose a model of oscillation of Davydov solitons in three wells. It can be used as a mathematical and physical frame in simulation of circle of some nonlinear oscillation of excitations via acupuncture system. The calculation shows that this sort of oscillation is possible if the initial rate of average occupational number of the quasi-particles in the wells is not equal to zero. One of oscillations arising relies on the initial rate of average occupational number of quasi-particles to be equal with each other within three wells. Then, the oscillation is not a kind of Josephson oscillation and has complicated frequency distributions. However, the total behavior of oscillation played is similar to three big solitons concentrated in three wells. In this sense, this model generally reveals a sort of oscillation mechanism of the acupuncture system how to work in the body, which allows us to understand the oscillation that may be one of fundamental natures in the acupuncture system.

## Keywords

Nonlinear Oscillation, Davydov Soliton, Acupuncture System

## 1. Introduction

Bio-excitations interaction with living organization of our body appears many nonlinear effects, which is a significant issue studied in biophysics, bio-photonics, or even life sciences [1]. In early 1973, Davydov has proposed protein molecules excited "solitary" model of the energy transport [2]-[4]. According to his theory, three spiral microvibration and lattice distortion of amide-I exciton in a protein molecule produce collective excitations to form a soliton, along the helix propagation, so that the adenosine triphosphate (ATP) molecules are hydrolyzed to produce energy from one place to another place. This can be found in the experiment that soliton resonance light decomposes into excitons and local deformation, corresponding to a new band in 1650 $cm^{-1}$, with amide-I exciton infrared absorption spectra observed on the 1666 $cm^{-1}$ line, which demonstrates that there is a red shift of 16 cm$^{-1}$ corresponding to the formation of just soliton bound energy. However, Davydov soliton seems to appear that short time span is serious obstacle to explain why it is a basic unit of energy and information transmission in bio-systems. For improving this weakness of the model, many scholars proposed modified models [5]. After that, Pang Xiaofeng improved and developed the Davydov soliton model with longer life span and established a sort of biological soliton transmission theory based on his nonlinear quantum theory [6], by which Pang Xiaofeng showed that the revised Davydov solitons could play a basic metabolism role in energy and information transmission of bio-systems including human body.

In this work, we propose a model of oscillation of the Davydov solitons, via an acupuncture system (so called medial vein), in three wells of potentials to simulate the Davydov solitons condensation in three areas of living system. The author hopes this studying is constructive to allow us understand that there may exist three condensations of solitons restricted in three areas in human body, which may provide a possible mechanism to explain working nature of three acupuncture points, *i.e.* upper, middle and lower Dantian as described by Chinese medicine and acupuncture theory [7].

## 2. Model of Oscillation

Let us consider a Davydov bio-vibration system in our body which is described by a nonlinear Schrödinger equation [8]

$$i\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t)=\left[-\frac{\nabla^2}{2m}+V_{ext}(\boldsymbol{r})+g\left|\Psi(\boldsymbol{r},t)\right|^2\right]\Psi(\boldsymbol{r},t), \tag{1}$$

where $g$ is a coupling number, the wave function $\Psi(\boldsymbol{r},t)$ is supposed to be divided into three parts to oscillate (via a so called medial vein) among three potential wells. Here, the medial vein is one of parts of the acupuncture system, which passes through three acupuncture points described by the upper, middle and lower Dantian [7]. Then, $\Psi(\boldsymbol{r},t)$ is expanded as

$$\Psi(\boldsymbol{r},t)=a(t)\varphi_1(\boldsymbol{r})+b(t)\varphi_2(\boldsymbol{r})+c(t)\varphi_3(\boldsymbol{r}), \tag{2}$$

where $\varphi_1(\boldsymbol{r})$, $\varphi_2(\boldsymbol{r})$ and $\varphi_3(\boldsymbol{r})$ is a basic wave function in three wells, respectively, while $a(t)$, $b(t)$ and $c(t)$ are correspondingly probability amplitude, with

$$|a(t)|^2+|b(t)|^2+|c(t)|^2=1. \tag{3}$$

By instead Equation (2) into Equation (1), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}a(t)\\b(t)\\c(t)\end{pmatrix}=\begin{pmatrix}\gamma+C|a(t)|^2 & -\frac{v}{2} & 0\\ -\frac{v}{2} & C|b(t)|^2 & -\frac{v}{2}\\ 0 & -\frac{v}{2} & -\gamma+C|c(t)|^2\end{pmatrix}\begin{pmatrix}a(t)\\b(t)\\c(t)\end{pmatrix}, \tag{4}$$

which can describe oscillation of soliton in three wells, where the relevant parameters are defined by

$$\gamma_j = \int \left[ \frac{1}{2m} \left| \nabla \varphi_j(\mathbf{r}) \right|^2 + V_{ext}(\mathbf{r}) \left| \varphi_j(\mathbf{r}) \right|^2 \right] \mathrm{d}\mathbf{r},\ j = 1, 2, 3, \tag{5}$$

$$v = 3\int \left[ \frac{1}{2m} \nabla \varphi_1(\mathbf{r}) \nabla \varphi_2(\mathbf{r}) \nabla \varphi_3(\mathbf{r}) + V_{ext}(\mathbf{r}) \varphi_1(\mathbf{r}) \varphi_2(\mathbf{r}) \varphi_3(\mathbf{r}) \right], \tag{6}$$

and

$$\gamma = |\gamma_2 - \gamma_1| = |\gamma_2 - \gamma_3|. \tag{7}$$

## 3. Solution of the Equation

Then through observation, the solution for Equation (4) can be approximately constructed by

$$\begin{pmatrix} a(t) \\ b(t) \\ c(t) \end{pmatrix} = \begin{pmatrix} |f_1| a'(t) \\ |f_2| b'(t) \\ |f_3| c'(t) \end{pmatrix}. \tag{8}$$

Then replacing Equation (8) into Equation (4), we can get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} |f_1| a'(t) \\ |f_2| b'(t) \\ |f_3| c'(t) \end{pmatrix} &= \begin{pmatrix} |f_1| \frac{\mathrm{d}a'(t)}{\mathrm{d}t} + a'(t) \frac{\mathrm{d}|f_1|}{\mathrm{d}t} \\ |f_2| \frac{\mathrm{d}b'(t)}{\mathrm{d}t} + b'(t) \frac{\mathrm{d}|f_2|}{\mathrm{d}t} \\ |f_3| \frac{\mathrm{d}c'(t)}{\mathrm{d}t} + c'(t) \frac{\mathrm{d}|f_3|}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} |f_1| \frac{\mathrm{d}a'(t)}{\mathrm{d}t} \\ |f_2| \frac{\mathrm{d}b'(t)}{\mathrm{d}t} \\ |f_3| \frac{\mathrm{d}c'(t)}{\mathrm{d}t} \end{pmatrix} + \begin{pmatrix} a'(t) \frac{\mathrm{d}|f_1|}{\mathrm{d}t} \\ b'(t) \frac{\mathrm{d}|f_2|}{\mathrm{d}t} \\ c'(t) \frac{\mathrm{d}|f_3|}{\mathrm{d}t} \end{pmatrix} \\ &= \left[ \begin{pmatrix} \gamma & -\frac{v}{2} & 0 \\ -\frac{v}{2} & 0 & -\frac{v}{2} \\ 0 & -\frac{v}{2} & -\gamma \end{pmatrix} + \begin{pmatrix} C|f_1 a'(t)|^2 & 0 & 0 \\ 0 & C|f_2 b'(t)|^2 & 0 \\ 0 & 0 & -\gamma + C|f_3 c'(t)|^2 \end{pmatrix} \right] \begin{pmatrix} |f_1| a'(t) \\ |f_2| b'(t) \\ |f_3| c'(t) \end{pmatrix}. \end{aligned} \tag{9}$$

This gives

$$\begin{pmatrix} |f_1| \frac{\mathrm{d}a'(t)}{\mathrm{d}t} \\ |f_2| \frac{\mathrm{d}b'(t)}{\mathrm{d}t} \\ |f_3| \frac{\mathrm{d}c'(t)}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{v}{2} & 0 \\ -\frac{v}{2} & 0 & -\frac{v}{2} \\ 0 & -\frac{v}{2} & -\gamma \end{pmatrix} \begin{pmatrix} |f_1| a'(t) \\ |f_2| b'(t) \\ |f_3| c'(t) \end{pmatrix}, \tag{10}$$

and

$$\begin{pmatrix} a'(t) \frac{\mathrm{d}|f_1|}{\mathrm{d}t} \\ b'(t) \frac{\mathrm{d}|f_2|}{\mathrm{d}t} \\ c'(t) \frac{\mathrm{d}|f_3|}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} C|f_1 a'(t)|^2 & 0 & 0 \\ 0 & C|f_2 b'(t)|^2 & 0 \\ 0 & 0 & C|f_3 c'(t)|^2 \end{pmatrix} \begin{pmatrix} |f_1| a'(t) \\ |f_2| b'(t) \\ |f_3| c'(t) \end{pmatrix}. \tag{11}$$

The first Equation (10) can be simplified as

$$\begin{pmatrix} \frac{\mathrm{d}a'(t)}{\mathrm{d}t} \\ \frac{\mathrm{d}b'(t)}{\mathrm{d}t} \\ \frac{\mathrm{d}c'(t)}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{v}{2}\lambda & 0 \\ -\frac{v}{2\lambda} & 0 & -\frac{v}{2}\eta \\ 0 & -\frac{v}{2\eta} & -\gamma \end{pmatrix} \begin{pmatrix} a'(t) \\ b'(t) \\ c'(t) \end{pmatrix}, \tag{12}$$

where the physical meaning of $\lambda$ and $\eta$ are approximately proportional to the rate of average occupational number of quasi-particles in the wells,

$$\lambda = \left|\frac{f_2}{f_1}\right| \sim \frac{\langle n_2 \rangle}{\langle n_1 \rangle}, \eta = \left|\frac{f_3}{f_2}\right| \sim \frac{\langle n_3 \rangle}{\langle n_2 \rangle}, \tag{13}$$

here $\langle n_1 \rangle$, $\langle n_2 \rangle$, $\langle n_3 \rangle$ is the average occupying quasi-particle number in three potential wells, respectively. Thus, the solution of Equation (12) is approximately gotten by

$$\begin{aligned} a'(t) &= -\frac{1}{v\sqrt{v^2+2\gamma^2}} \Bigg( 2C_1\gamma\eta\sqrt{v^2+2\gamma^2} + \left(2\gamma\sqrt{v^2+2\gamma^2} + \sqrt{2}v^2 + 2\sqrt{2}\gamma^2\right) C_2\eta \mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t} \\ &\quad + \left(2\gamma\sqrt{v^2+2\gamma^2} - \sqrt{2}v^2 - 2\sqrt{2}\gamma^2\right) C_3\eta \mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t} \Bigg) \\ &= A_1 + A_2\mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t} + A_3\mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}, \\ b'(t) &= \left(5C_2 + \frac{2\sqrt{2}}{v^2}C_2\gamma\sqrt{v^2+2\gamma^2}\right)\lambda\eta \mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t} + \left(5C_3 - C_1 - \frac{2\sqrt{2}}{v^2}C_3\gamma\sqrt{v^2+2\gamma^2}\right)\lambda\eta \mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t} \\ &= B_1\mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t} + B_2\mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}, \\ c'(t) &= C_1 + C_2\mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t} + C_3\mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}, \end{aligned} \tag{14}$$

where $C_1$, $C_2$, and $C_3$ are three integral constants which can be determined by the initial conditions $a'(0)$, $b'(0)$, and $c'(0)$. For instance, if the average occupational number is equal with each other in three wells then $C_1$, $C_2$ and $C_3$ are determined by

$$\begin{aligned} C_1 &= -\frac{1}{6v^2}\left(b'(0)v^2 - c'(0)\left(5v^2 - 4\gamma^2\right) + 2a'(0)v\gamma\right), \\ C_2 &= \frac{1}{12\sqrt{2}\left(v^4 + 2v^2\gamma^2\right)}\Big(c'(0)\left(8\sqrt{2}\gamma^4 + \sqrt{2}v^4 + 6\sqrt{2}v^2\gamma^2 - 12v^2\gamma\sqrt{2\gamma^2+v^2}\right) \\ &\quad - a'(0)2\left(3v^3\sqrt{2\gamma^2+v^2} - 2\sqrt{2}v\gamma^3 - \sqrt{2}v^3\gamma\right) + b'(0)\sqrt{2}v^2\left(2\gamma^2+v^2\right)\Big), \\ C_3 &= \frac{1}{12\sqrt{2}\left(v^4 + 2v^2\gamma^2\right)}\Big(c'(0)\left(8\sqrt{2}\gamma^4 + \sqrt{2}v^4 + 6\sqrt{2}v^2\gamma^2 + 12cv^2\gamma\sqrt{2\gamma^2+v^2}\right) \\ &\quad + a'(0)2\left(3v^3\sqrt{2\gamma^2+v^2} + 2\sqrt{2}v\gamma^3 + \sqrt{2}v^3\gamma\right) + b'(0)\sqrt{2}v^2\left(2\gamma^2+v^2\right)\Big) \end{aligned} \tag{15}$$

here the conditions $12\sqrt{2}\left(v^4+2v^2\gamma^2\right)\neq 0$ and $\left(8\sqrt{2}\gamma^4+\sqrt{2}v^4+6\sqrt{2}v^2\gamma^2+12cv^2\gamma\sqrt{2\gamma^2+v^2}\right)\neq 0$ have to be considered for the above Equation (15) to have mathematical meaning.

On the other hand, for solving the second Equation (11) it is obviously to have

$$
\begin{aligned}
|f_1| &= \left(-\int 2C\left|a'(t)\right|^2 \mathrm{d}t\right)^{-\frac{1}{2}} \\
&= \left(-\int 2C\left|A_1+A_2\mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}+A_3\mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}\right|^2 \mathrm{d}t\right)^{-\frac{1}{2}} \\
&= \left[-2CA_1^2t-\frac{\sqrt{2}CA_2^2}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{\sqrt{2}\sqrt{v^2+2\gamma^2}t}+\frac{\sqrt{2}CA_3^2}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{-\sqrt{2}\sqrt{v^2+2\gamma^2}t}\right. \\
&\quad \left.-\frac{4\sqrt{2}CA_1A_2}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}+\frac{4\sqrt{2}CA_1A_3}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}-2CA_2A_3t\right]^{-\frac{1}{2}},
\end{aligned}
\tag{16}
$$

$$
\begin{aligned}
|f_2| &= \left(-\int 2C\left|b'(t)\right|^2 \mathrm{d}t\right)^{-\frac{1}{2}} \\
&= \left(-\int 2C\left|B_1\mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}+B_2\mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}\right|^2 \mathrm{d}t\right)^{-\frac{1}{2}} \\
&= \left(\frac{-\sqrt{2}CB_1^2}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{\sqrt{2}\sqrt{v^2+2\gamma^2}t}+\frac{\sqrt{2}CB_2^2}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{-\sqrt{2}\sqrt{v^2+2\gamma^2}t}-4CB_1B_2t\right)^{-\frac{1}{2}},
\end{aligned}
\tag{17}
$$

and

$$
\begin{aligned}
|f_3| &= \left(-\int 2C\left|c'(t)\right|^2 \mathrm{d}t\right)^{-\frac{1}{2}} \\
&= \left(-\int 2C\left|C_1+C_2\mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}+C_3\mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}\right|^2 \mathrm{d}t\right)^{-\frac{1}{2}} \\
&= \left(-2CC_1^2t-\frac{\sqrt{2}CC_2^2}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{\sqrt{2}\sqrt{v^2+2\gamma^2}t}+\frac{CC_3^2}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{-\sqrt{2}\sqrt{v^2+2\gamma^2}t}\right. \\
&\quad \left.-\frac{4\sqrt{2}CC_1C_2}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}+\frac{4\sqrt{2}CC_1C_3}{\sqrt{v^2+2\gamma^2}}\mathrm{e}^{-\frac{\sqrt{2}}{2}\sqrt{v^2+2\gamma^2}t}-2CC_2C_3t\right)^{-\frac{1}{2}}.
\end{aligned}
\tag{18}
$$

Therefore, we eventually obtain the solution $a(t), b(t)$ and $c(t)$ by combining Equations (16)-(18) with Equation (14), namely

$$
\begin{pmatrix} a(t) \\ b(t) \\ c(t) \end{pmatrix} = \begin{pmatrix} |f_1|a'(t) \\ |f_2|b'(t) \\ |f_3|c'(t) \end{pmatrix}. \tag{19}
$$

This shows that the oscillation posesses complicated frequency distributions and therefore it is not a Josephson oscillation as $a'(t)$, $b'(t)$, and $c'(t)$ with the period $2\pi\Big/\sqrt{\frac{v^2}{2}+\gamma^2}$, for $C=0$ [8]-[10]. Secondly, the initial distribution of the average ocpying quasi-particle numbers for the oscillation $C\neq 0$ in three potential wells should be not equal to 0, for instance, if they are equal with each other in three wells, *i.e.*

$$\lambda=\left|\frac{f_2}{f_1}\right|\sim\frac{\langle n_2\rangle}{\langle n_1\rangle}=1,\ \eta=\left|\frac{f_3}{f_2}\right|\sim\frac{\langle n_3\rangle}{\langle n_2\rangle}=1, \tag{20}$$

then the total oscillation will take part in three big solitons confined in three wells, respectively, which are described by Equation (19).

## 4. Conclusion and Remarks

In conclusions, the above model can be used in simulation of circle of some nonlinear excitations moving in the acupuncture system in advanced level, such as meaning of qi (here can be understood that the Davydov solitons is a kind of qi) moving in the medial vein and oscillation in three Dantians. The calculation shows that this sort of oscillation is possible if the initial rate of average occupational number of the quasi-particles in the wells is not equal to zero. One of simple oscillations arising relies on the initial rate of average occupational number of quasi-particles to be equal with each other within three wells, then the oscillation is not a kind of Josephson oscillation and has complicated frequency distributions, however the total behavior of oscillation played is similar to three big solitons concentrated in three wells. In the same way, the model of oscillation can be extended to oscillation over three wells system, such as oscillation in seven wells, to mimic working function of complicated medial vein + circle system which is also described by the theory of ancient Indian Yoga. In this sense, this model generally reveals a sort of oscillation mechanism of acupuncture system to working in our body, and explains clearly what means oscillation of qi, *i.e.* oscillation of Davydov solitons. This allows us to propose an assumption: there is no clear bio-structure (in the level of anthropotomy) and only has functional structure for acupuncture system. The oscillation of various excitatoins [11] falls into the acupuncture system, by which the information and energy can be transmitted to many places of the body and returning message is also the feedback for regulating the body. This is an oscillation which may be one of fundamental natures of our acupuncture system. We hope this studying could expose certain physical essences of the circle of qi in three Dantian via the medial vein for helping practitioners to increase their level.

## References


[1] Shen, X. and van Wijk, R. (2006) Biophotonics: Optical Science and Engineering for the 21st Century. Springer, Berlin.

[2] Davydov, A.S. (1979) *Physica Scripta*, **20**, 387-396. http://dx.doi.org/10.1088/0031-8949/20/3-4/013

[3] Davydov, A.S. and Kislukha, N.I. (1973) *Physica Status Solidi* (*b*), **59**, 465-470. http://dx.doi.org/10.1002/pssb.2220590212

[4] Davydov, A.S. (1973) *Journal of Theoretical Biology*, **38**, 559-569. http://dx.doi.org/10.1016/0022-5193(73)90256-7

[5] Pang, X.F. and Feng, Y.-P. (2005) Quantum Mechanics in Nonlinear Systems. World Scientific Publishing Co. Pte. Ltd., Singapore City.

[6] Pang, X.-F. (2000) *Physical Review E*, **62**, 6989-6998. http://dx.doi.org/10.1103/PhysRevE.62.6989

[7] Oconnor (1992) Acupuncture: A Comprehensive Text. Shanghai College of Traditional Medicine, Shanghai.

[8] Liu, J. (2009) Dynamics for the Bose-Einstein Condensation. Science Press, Beijing (in Chinese).

[9] Smerzi, A., Fantoni, S., Giovanazzi, S., *et al.* (1997) *Physical Review Letters*, **79**, 4950-4953. http://dx.doi.org/10.1103/PhysRevLett.79.4950

[10] Albiez, M., Gati, R., Fölling, J., *et al.* (2005) *Physical Review Letters*, **95**, 010402-5.

[11] Qiao, B., Song, K.Z. and Ruda, H.E. (2012) *Journal of Modern Physics*, **3**, 1907-1913. http://dx.doi.org/10.4236/jmp.2012.312240

Journal of Modern Physics, 2016, 7, 1818-1828
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# New Basic Theory of Gravity

**Hubert J. Veringa**

Department Mechanical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands
Email: Veringa48@planet.nl

**How to cite this paper:** Veringa, H.J. (2016) New Basic Theory of Gravity. *Journal of Modern Physics*, **7**, 1818-1828.
http://dx.doi.org/10.4236/jmp.2016.713162

**Received:** September 5, 2016
**Accepted:** September 27, 2016
**Published:** September 30, 2016



## Abstract

Although Newton's law of gravity already exists for centuries, and its validity is beyond any doubt, we are still lacking a basic theory to explain the specific features of this law. The general belief is that any suitable theory should include, or will be a merger of, classical quantum theory and general relativity, but until now no acceptable mathematical model taking both aspects into account has proposed. The present letter is written to present a new scheme of analysis for the mutual interaction between particles that have some exchange with respect to time and space. It is found that the right form of Newton's gravity law emerges by consequently working through the existing schemes of both quantum mechanics and the basic equations of relativity theory as expressed by the Dirac equation.

## Keywords

Gravity, Quantum Physics, General Relativity

## 1. Introduction

Newton's law of gravity is the cornerstone of many disciplines in space technology, astronomy and cosmology, but so far no acceptable theory to explain its peculiarities is existing. It is broadly accepted that any suitable theory should include, or will be a merger of, classical quantum theory and general relativity [1] [2]. The main problem is, however, that the basic equations of quantum mechanics and relativity are not in correspondence. This fact is most easily seen that, when the common factor in the Dirac equation and the Schrödinger equation are used to combine both, a contradiction shows up. The reason for this contradiction is that relativity requires co-variance throughout where the Schrödinger equation is not co-variant. There are examples where this contradiction is circumvented like in the theory of the magnetic moment of particles, leading to the spins. In the present paper, a new scheme of analysis is presented for the mutual interaction between particles that have some exchange with respect to time and space. The specific requirements on invariance and co-variance of operators and quantities will be carefully taken care of and are finally found to be of great importance for the result. This pair formation is described quantum mechanically, either starting from the classical Schrödinger equation or the relativistic Dirac equation. This latter is formulated in a quantum mechanical setting. Both result in the same wave function describing pairs of particles. Since this wave function represents a pair potential, a relativistic mass can be attributed to it which is used in the Dirac equation to derive an interaction field between the members that form the ensemble. It is found that the right form of Newton's gravity law emerges by consequently working through the proposed schemes of both quantum mechanics and the basic equations of relativity theory as expressed by the Dirac equation[1].

## 2. Forming of Pairs

Starting point is the assumption that there are two independent particles indicated by the masses $m_i$ and $m_j$ which are described by the normal Schrödinger equation. In the present treatment the kinetic energy is taken into account and they experience some force reflected by the potential $V_i$ and $V_j$. Spherical symmetry is next adopted and the only boundary condition is that the wave function is zero at infinity. An observer at $m_i$ at a distance $r_{ij}$ from particle $m_j$ and another on $m_j$ at $r_{ji}$ from particle $m_i$ will see that the total wave equation is defined as follows [5]:

$$\widehat{H_{ij}}\Psi_{ij,t} = i\hbar\frac{\partial}{\partial t}\Psi_{ij,t} = -\left(\frac{1}{r_{ij}^2}\frac{\partial}{\partial r_{ij}}r_{ij}^2\frac{\partial}{\partial r_{ij}} + \frac{\hbar^2}{2m_j}\frac{1}{r_{ji}^2}\frac{\partial}{\partial r_{ji}}r_{ji}^2\frac{\partial}{\partial r_{ji}}\right)\Psi_{ij,t} + \left(V_i + V_j\right)\Psi_{ij,t}. \quad (1)$$

where $\Psi_{ij,t}$ is the time and space dependent wave function. The time dependence can be removed by replacing the time dependent wave function $\Psi_{ij,t}$ by $\Psi_{ij}e^{iE_{ij}t/\hbar}$. Further define $V_i + V_j$ by $V_{ij}$ and we get:

$$\left(E_{ij} - V_{ij}\right)\Psi_{ij} + \frac{\hbar^2}{2m_i}\frac{1}{r_{ij}^2}\frac{\partial}{\partial r_{ij}}r_{ij}^2\frac{\partial}{\partial r_{ij}}\Psi_{ij} + \frac{\hbar^2}{2m_j}\frac{1}{r_{ji}^2}\frac{\partial}{\partial r_{ji}}r_{ji}^2\frac{\partial}{\partial r_{ji}}\Psi_{ij} = 0. \quad (2)$$

To simplify the equation replace $E_{ij} - V_{ij}$ by $\varepsilon_{ij}$ to propose a solution that is valid in areas where the $V_{ij}$ is not of great influence anymore as follows:

$$\Psi_{ij} = \left(\frac{\alpha_{ij}}{r_{ij}} + \frac{\alpha_{ji}}{r_{ji}}\right)e^{i\beta_{ij}r_{ij} + i\beta_{ji}r_{ji}}, \quad (3)$$

where $\alpha_{ij}$ and $\beta_{ij}$ are constants independent of space coordinates and time. This solution means that we consider the wave function outside the surroundings where the potential energy with all its peculiarities has a very minor effect on the shape of the wave function. The only interaction that can play a role will then be based solely on gravitational interaction. By substituting the solution in Equation (3) the following relation is found:

[1]The analysis is based on standard arguments of quantum theory and specific and general relativity theory. Specific reading on the separate subjects in monographs which are easily accessible to appreciate the foregoing analysis can be found in [3] [4] together.

$$-\frac{\hbar^2 i}{r_{ij}r_{ji}}\left(\frac{\alpha_{ij}\beta_{ji}}{m_j}+\frac{\alpha_{ji}\beta_{ij}}{m_i}\right)\mathrm{e}^{i\beta_{ij}r_{ij}+i\beta_{ji}r_{ji}}-\frac{\hbar^2}{2}\left(\frac{\beta_{ij}^2}{m_i}+\frac{\beta_{ji}^2}{m_j}\right)\left(\frac{\alpha_{ij}}{r_{ij}}+\frac{\alpha_{ji}}{r_{ji}}\right)\mathrm{e}^{i\beta_{ij}r_{ij}+i\beta_{ji}r_{ji}}$$
$$+\varepsilon_{ij}\left(\frac{\alpha_{ij}}{r_{ij}}+\frac{\alpha_{ji}}{r_{ji}}\right)\mathrm{e}^{i\beta_{ij}r_{ij}+i\beta_{ji}r_{ji}}=0. \tag{4}$$

The first term at the left hand side is to be set to zero so that in a pair-wise process $\alpha_{ij}\beta_{ji}/m_j+\alpha_{ji}\beta_{ij}/m_i=0$ and $\beta_{ij}^2\hbar^2/2m_i+\beta_{ji}^2\hbar^2/2m_j=\varepsilon_{ij}$.

At the moment not much is known about the *α's*, but one requirement to be imposed on the wave function is that it represents a pair of particles. For the time being it can be said that:

a) The *α's* cannot depend on the running variables in the wave equation: $r_{ij}$ or $t$. It will be a constant that can only depend on fundamental nature constants and the particle masses.

b) It should make no difference for the outside world how one member sees its partner or whether and how we see the two members of the pair. It means that most likely we can say: $\alpha_{ij}=f(m_i)f(m_j)$.

c) There is no pair if either $m_i$ or $m_j$ equals zero so that $f(m_i)=0$ for $m_i=0$.

Later it will be found that, for the sake of symmetry in the mutual gravitational interaction, the two *α's* should be equal. It also means that the *β's* have opposite signs and fixed values and by taking the *α's* equal we make their values independent of the masses and the energies of the members of the pair. The $\varepsilon_{ij}$ could have been split into two separate quantities as $\varepsilon_{ij}$ and $\varepsilon_{ji}$ to dedicate the $\beta_{ij}^2$ and $\beta_{ji}^2$-values to the separate energies of the two particles. It is also interesting to notice that the solution of the wave equation for the pair looks different from a solution for a single particle:

$$\Psi_i=\left(\frac{\alpha_i}{r_i}\right)\mathrm{e}^{i\beta_i r_i}. \tag{5}$$

For instance if we take a look at the $r_i$ dependence in the solution (3) we see that there is an extra $r_2$ dependent factor in the exponential term. This latter term is insufficient to make such a solution applicable for the operator working on $r_2$. For it to be sufficient we need the total pre-exponential factor as given in Equation (3).

Another approach is taking the Dirac equation as the starting point. In this way, we guarantee full co-variance throughout the entire analysis. The Dirac equation reads [6]:

$$E^2-p^2c^2=m_0^2c^4$$

or expressed alternatively:

$$E^2/m_0^2c^4-p^2c^2/m_0^2c^4=1,$$

and translated into quantum mechanical language for an ensemble of two particles:

$$\left(E_{ij}^2/m_{i,0}^2c^4-E_{ji}^2/m_{j,0}^2c^4\right)\Psi_{ij}-\left(\left(\widehat{p_{ij}}\right)^2/m_{i,0}^2c^2-\left(\widehat{p_{ji}}\right)^2/m_{j,0}^2c^2\right)\Psi_{ij}=0. \tag{6}$$

where $\widehat{p_{ij}}$ is the momentum operator to be written out in spherical coordinates as in Equation (1) and $m_{i,o}$ the rest mass of the particle *i* in the ensemble *ij*. Also in this case it immediately can be seen that, with the solution of the form as in Equation (3), the same interpretation as before can be given.

For the $\alpha$'s it means that:

d) If the energy of the ensemble should be related to the masses of each member, which definitely is the case outside the area where the potential energy is not of any importance, it follows that: $\beta_{ij}^2\hbar^2/2m_i + \beta_{ji}^2\hbar^2/2m_j = \varepsilon_{ij} = \sigma(m_i + m_j)$ and a similar equation in the case that the Dirac equation is taken as starting point. It means, following the other boundary condition $\alpha_{ij}\beta_{ji}/m_j + \alpha_{ji}\beta_{ij}/m_i = 0$, that $\alpha_{ij} = \alpha_{ji}$. This is in accordance with the argument given in point b). For the transparency of the analysis, we will, however, not yet take into account that the $\alpha$'s are equal. Leaving the statement of the equality of the $\alpha_{ij}$ values for later has an interesting, causality related, consequence on the symmetry of the gravitational interaction between particles.

It appears that with the simple assumption of having a pair of particles, and taking a wave equation with spherical symmetry, a solution is obtained that, apparently, couples the particles into pairs. It is surprising that the procedure only works well with sets of two particles.

The wave function as derived gives the presence of an entity for which it is derived. In this case it is the pair potential so that a mass can be dedicated to this potential defined as $\Psi_{ij}^* m_o^2 \Psi_{ij}$ and which becomes equal to:

$$\Psi_{ij}^* m_0^2 \Psi_{ij} = \left(\frac{\alpha_{ij}}{r_{ij}} + \frac{\alpha_{ji}}{r_{ji}}\right)^2. \tag{7}$$

As said before, in this expression the $m_0^2$ which occurs in the Dirac equation can be identified as a quantity that represents the presence of a pair of particles. It is related to the mass of the pair since the product of the complex conjugated wave function and the wave function with the appropriate operator, in this case the $m_0^2$, gives the expectation value of the operator. The $\alpha$-values in this last equation accommodate the influence of this $m_0^2$ but, as it follows from $\alpha_{ij}\beta_{ji}/m_j + \alpha_{ji}\beta_{ij}/m_i = 0$ with $\alpha_{ij} = \alpha_{ji}$ that there is some freedom in choosing its dependence on relativistic parameters such that the right hand side of Equation (7) becomes an invariant as it should be.

## 3. Relativistic Interaction

Now, as a next step, the pair is considered as essentially one entity and the problem can be analysed in the relativistic four dimensional space where the Dirac equation is the appropriate starting point [1] [6]:

$$E^2 - p^2c^2 = m_0^2c^4 \quad \text{or} \quad -p^2c^2 = m_0^2c^4 - E^2.$$

Again we will have to translate this equation into the appropriate quantum mechanical language for pairs as one entity and therefore make the following transformations:

$$-p^2c^2\varphi_{ij,t}\varphi_{ji,t} = \left(m_0^2c^4 - E^2\right)\varphi_{ij,t}\varphi_{ji,t},$$

and

$$
\begin{aligned}
p^2 &= \left(\widehat{p_{ij}} + \widehat{p_{ji}}\right)^2 \\
&= -\hbar^2 \left( \frac{1}{r_{ij}^2}\frac{\partial}{\partial r_{ij}} r_{ij}^2 \frac{\partial}{\partial r_{ij}} + \frac{1}{r_{ji}^2}\frac{\partial}{\partial r_{ji}} r_{ji}^2 \frac{\partial}{\partial r_{ji}} + \frac{1}{r_{ij}^2}\frac{\partial}{\partial r_{ij}} r_{ij}^2 \frac{\partial}{\partial r_{ji}} + \frac{1}{r_{ji}^2}\frac{\partial}{\partial r_{ji}} r_{ji}^2 \frac{\partial}{\partial r_{ij}} \right).
\end{aligned}
$$

The last expression is, as different from earlier, a mixed sum of the momenta. This representation is a consequence of the fact that the particles have been treated only in pairs and that spherical symmetry remains to be adopted. Referring to **Figure 1** the total relativistic Dirac equation for an undefined number of pairs ($ij$) is set up. There are $\mathcal{N}$ particles which make a total of $\mathcal{N} = N!/(2(N-2)!)$ pairs, each of which are described by a wave function as a solution of the initial Schrödinger equation. As before the $\alpha$-values accommodate all necessary multiplication constants. Adding up for all pairs and treating them as mutually independent and taking into account the basic rules of quantum mechanics, lead to:

$$
\begin{aligned}
& c^2\hbar^2 \sum\nolimits_{ij} \left( \frac{1}{r_{ij}^2}\frac{\partial}{\partial r_{ij}} r_{ij}^2 \frac{\partial}{\partial r_{ij}} + \frac{1}{r_{ji}^2}\frac{\partial}{\partial r_{ji}} r_{ji}^2 \frac{\partial}{\partial r_{ji}} + \frac{1}{r_{ij}^2}\frac{\partial}{\partial r_{ij}} r_{ij}^2 \frac{\partial}{\partial r_{ji}} + \frac{1}{r_{ji}^2}\frac{\partial}{\partial r_{ji}} r_{ji}^2 \frac{\partial}{\partial r_{ij}} \right) \times \prod\nolimits_{ij} \varphi_{ji,t}\varphi_{ij,t} \\
&= \sum\nolimits_{ij} \left( \frac{\alpha_{ij}^2}{r_{ij}^2} + 2\frac{\alpha_{ij}}{r_{ij}}\frac{\alpha_{ji}}{r_{ji}} + \frac{\alpha_{ji}^2}{r_{ji}^2} \right) \times \prod\nolimits_{ij} \varphi_{ji,t}\varphi_{ij,t} - \sum\nolimits_{ij} \left( \widehat{E_i} + \widehat{E_j} \right)^2 \times \prod\nolimits_{ij} \varphi_{ij,t}\varphi_{ji,t}.
\end{aligned}
\tag{8}
$$

As before the time dependences can be removed by setting:

$$
\varphi_{ij,t}\varphi_{ji,t} = \varphi_{ij}\varphi_{ji}\mathrm{e}^{i\left(E_{ij}+E_{ji}\right)t/\hbar}, \tag{9}
$$

so that:

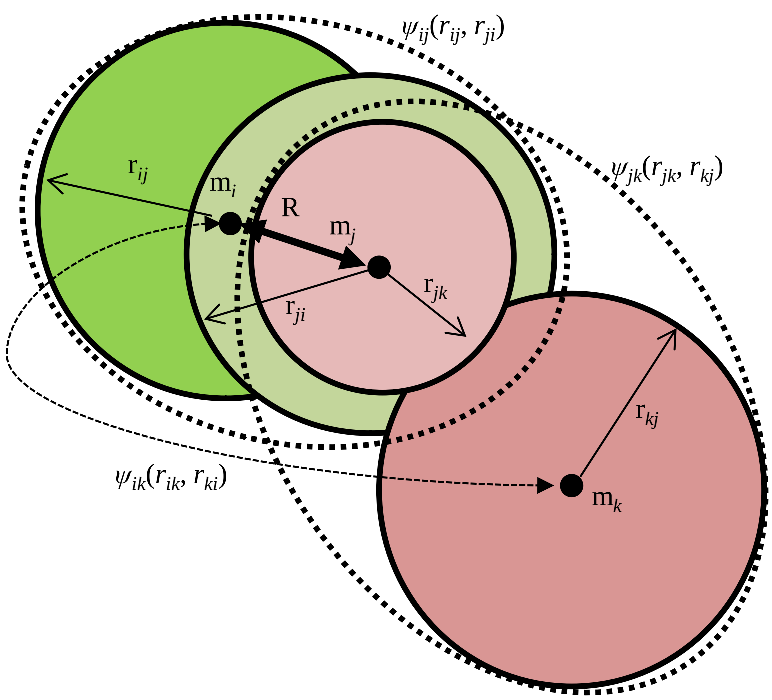



**Figure 1.** Forming and describing of $\mathcal{N} = N!/2(N-2)!$ pairs.

$$\sum_{ij}\left(\widehat{E}_i+\widehat{E}_j\right)^2\varphi_{ij,t}\varphi_{ji,t}=\sum_{ij}\left(E_{ij}+E_{ji}\right)^2\varphi_{ij}\varphi_{ji}. \tag{10}$$

If all $\alpha's$ would have been equal to zero, a propagating wave $\varphi_{ij,t}\varphi_{ji,t}$ extending in the radial direction with the light velocity would have resulted. Non zero values of $\alpha$ reduce this speed and, as a consequence, give mass to the field $\varphi_{ij,t}\varphi_{ji,t}$. The proposed solution will be:

$$\varphi_{ij}=\gamma_{ij}r_{ij}^{\alpha_{ij}/\hbar c}, \tag{11}$$

which is inserted into:

$$\begin{aligned}&\sum_{ij}\left(E_{ij}+E_{ji}\right)^2\times\prod_{ij}\varphi_{ij}\varphi_{ji}\\&+c^2\hbar^2\sum_{ij}\left(\frac{1}{r_{ij}^2}\frac{\partial}{\partial r_{ij}}r_{ij}^2\frac{\partial}{\partial r_{ij}}+\frac{1}{r_{ji}^2}\frac{\partial}{\partial r_{ji}}r_{ji}^2\frac{\partial}{\partial r_{ji}}+\frac{1}{r_{ij}^2}\frac{\partial}{\partial r_{ij}}r_{ij}^2\frac{\partial}{\partial r_{ji}}+\frac{1}{r_{ji}^2}\frac{\partial}{\partial r_{ji}}r_{ji}^2\frac{\partial}{\partial r_{ij}}\right)\times\prod_{ij}\varphi_{ij}\varphi_{ji}\\&=\sum_{ij}\left(\frac{\alpha_{ij}^2}{r_{ij}^2}+2\frac{\alpha_{ij}}{r_{ij}}\frac{\alpha_{ji}}{r_{ji}}+\frac{\alpha_{ji}^2}{r_{ji}^2}\right)\times\prod_{ij}\varphi_{ij}\varphi_{ji}.\end{aligned} \tag{12}$$

From the boundary condition that $\varphi_{ij}\left(r_{ji},\alpha_{ij}\right)=0$ for $r_{ji}$ to infinity a fifth condition on the $\alpha's$ can be derived:

e) $\alpha_{ij}$ is negative under all circumstances.

Putting all five conditions on $\alpha_{ij}$ together we can already conclude that the explicit expression for it is:

f) $\alpha_{ij}=\alpha_{ji}=-\sigma'\left(m_i m_j\right)^{2n}$ with n equal to 1, 2 etc.

Now, some algebra needs to be done, but in order to redistribute the various contributions it is easiest to start from the simplified equation:

$$\begin{aligned}&\left(E_{ij}^2+2E_{ij}E_{ji}+E_{ji}^2\right)\varphi_{ij}\varphi_{ji}+c^2\hbar^2\left(\frac{2}{r_{ij}}\varphi_{ij}\frac{\partial}{\partial r_{ji}}\varphi_{ji}+\frac{2}{r_{ji}}\varphi_{ji}\frac{\partial}{\partial r_{ij}}\varphi_{ij}+2\frac{\partial^2}{\partial r_{ij}\partial r_{ji}}\varphi_{ij}\varphi_{ji}\right)\\&+c\hbar\left(\frac{\alpha_{ij}}{r_{ij}^2}+\frac{\alpha_{ji}}{r_{ji}^2}\right)\varphi_{ji}\varphi_{ij}=2\frac{\alpha_{ij}}{r_{ij}}\frac{\alpha_{ji}}{r_{ji}}\varphi_{ji}\varphi_{ij}.\end{aligned} \tag{13}$$

with the solution proposed in Equation (11) it can immediately be seen that the sixth term in the first line is equal to the right hand side.

At this point a remark has to be made: removing the term $\alpha_{kl}^2/r_{kl}^2$ means that some basic interaction occurs between the gravitational field and the particle. Obviously, for this separate term, a Dirac equation can be formulated that shows that an entity with some relativistically derived mass operates and leaves behind a contribution to the interaction energy in the Equation (13). So already at this point there is direct interaction between the pair and the field around. Also removing the sixth term left together with the only remaining term at the right means that there is third interaction between the fields and the pair.

Taking all these interactions into account it is seen that all $\alpha$-terms in Equation (12) have disappeared. This has a profound meaning: in this model gravity is due to second order effects of peculiarities of the spherical symmetry in a relativistic setting. The effect is weak and operates over a long range so:

$$\left(E_{ij}^2 + 2E_{ij}E_{ji} + E_{ji}^2\right)\varphi_{ij}\varphi_{ji} + c^2\hbar^2\left(\frac{2}{r_{ij}}\varphi_{ij}\frac{\partial}{\partial r_{ji}}\varphi_{ji} + \frac{2}{r_{ji}}\varphi_{ji}\frac{\partial}{\partial r_{ij}}\varphi_{ij}\right)$$
$$+c\hbar\left(\frac{\alpha_{ij}}{r_{ij}^2} + \frac{\alpha_{ji}}{r_{ji}^2}\right)\varphi_{ji}\varphi_{ij} = 0. \tag{14}$$

The contributions can now be redistributed, but first multiply all terms by $r_{ij}r_{ji}$ and observe that the proposed solution is the only one that gives a sharp value for the quantity $E_{ij}r_{ij}$ and $E_{ji}r_{ji}$:

$$\left(E_{ij}^2 r_{ji}r_{ij} + E_{ij}E_{ji}r_{ji}r_{ij}\right)\varphi_{ij}\varphi_{ji} + 2c\hbar\alpha_{ij}\varphi_{ij}\varphi_{ji} + c\hbar\alpha_{ij}\frac{r_{ji}}{r_{ij}}\varphi_{ij}\varphi_{ji} = 0, \tag{15a}$$

$$\left(E_{ji}^2 r_{ji}r_{ij} + E_{ij}E_{ji}r_{ji}r_{ij}\right)\varphi_{ij}\varphi_{ji} + 2c\hbar\alpha_{ji}\varphi_{ij}\varphi_{ji} + c\hbar\alpha_{ji}\frac{r_{ij}}{r_{ji}}\varphi_{ij}\varphi_{ji} = 0. \tag{15b}$$

Cutting the Equation (14) into two separate ones as given in Equations (15a) and (15b) looks like arbitrary, as any cut between terms can be made, but if we now come back to the original suggestion as made in f), we a see that the gravitational interaction becomes symmetric. The gravitational energy of particle *i* is equal to the gravitational energy of particle *j*. It also reflects the point that a pair has to be seen as one entity. The observer cannot distinguish between the separate members of the pair.

It is also important to notice that the operators $\widehat{E_k}$ and $r_l$ commute. It means that "*Er*" is the quantity that has a sharp value, meaning that *E* has a sharp value if *r* is well defined.

## 4. Law of Gravity

Most important for finding out how the members of a pair see each other is to look at the Equations (15a) and (15b) by an observer on $m_i$ who sees the particle $m_j$ at a distance of $r_{ij}$ and an observer on particle $m_j$ looking at $m_i$ from a distance $r_{ji}$. Both see each other from the same distance $r_{ij} = r_{ji} = R$ and they already know that $\alpha_{ij} = \alpha_{ji} = -\alpha'$. There are no operators anymore in Equation (15), and they can conclude that $E_{ij} = E_{ji} = E$. This is an important conclusion. Obviously an electron and a proton forming a pair will have mutual interactions which are the same although their masses differ by some factor of about 1800 [7]. The result is a simple relation:

$$E^2R^2 = 3c\hbar\alpha'/2. \tag{16}$$

The boundary condition is that $\varphi_{ij}\varphi_{ji}$ goes to zero for r to infinity so that $\alpha' > 0$, and because both particles in the pair change their energy by the same amount. It follows for the two members of the ensemble together that:

$$2ER = \sqrt{6c\hbar\alpha'}, \tag{17}$$

and the gravitational force is given by: $-\partial 2E/\partial R = \text{constant}/R^2$.

All the work done to describe the total gravitational force, or rather the potential energy, has been based on the idea that all pairs that have been formed are acting independently so that we can add all the contributions of different masses constituting bodies in the real world without any interference.

Now it is important to see how pairs consisting of particles of different masses present themselves in $\alpha'$. As a consequence of Einstein's law that the rest energy of a particle is proportional to its mass which is also a direct consequence of the fact that the pairs which gives gravitational interaction are acting as single entities, and in view of Equation (17) we can only conclude that the exponent $n$ in condition f) is equal to one. As a consequence of the result in this equation the attractive force between two particles is proportional to the product of the two interacting masses. It also follows that due to gravitational interaction which carries energy, and for which a separate Dirac equation can be set up, some mass, although not much, is attributed to the pairs.

The final result is:

$$E = \sqrt{6\sigma' c\hbar} \cdot \left(m_i m_j\right) / R \tag{18}$$

Referring to Figure 2 where two masses $M_1$ and $M_2$ have particles numbered as $m_{1j}$ and $m_{2k}$ form $N_1 \times N_2$ pairs described by $\phi_{kl} = \varphi_{1k}\varphi_{2l}$ in which each $kl$-combination contributes separately to the interaction energy. Adding up all the interactions between

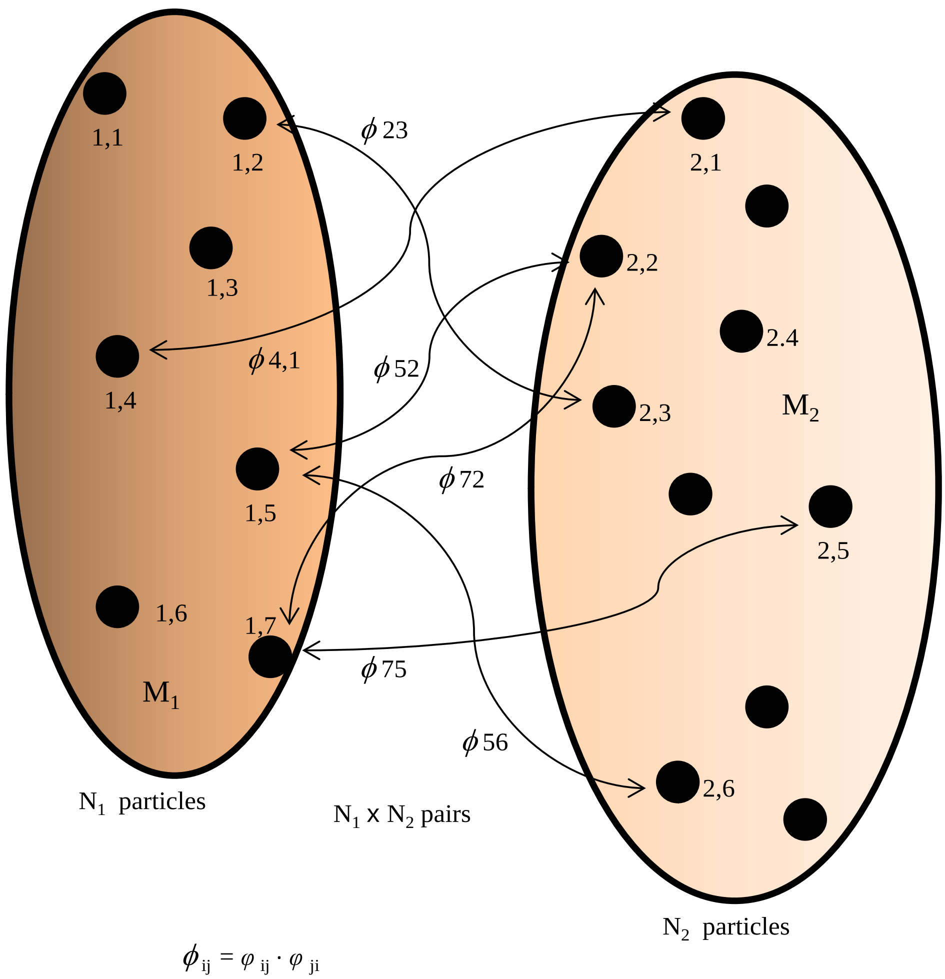



Figure 2. Interaction between masses.

particles, which in principle see each other at different distances which is a problem that has already been solved in the formulation of the classical theory of electrostatics [8].

Finally, Newton's gravitation law is obtained which reads: $div\boldsymbol{g} = 4\pi\rho G$ in which $\boldsymbol{g}$ is defined as a gravitational field around an entity constituting a space coordinates dependent mass density $\rho$. $G$ is the well known gravitational constant equal to: 6.673 × $10^{-11}$ $m^3 \cdot kg^{-1} \cdot sec^{-2}$.

In accordance with the theory of electrostatics the gravity law can also be given in vector representation for bodies $M_1$ and $M_2$ which have their centres of gravity at a separation of $R$:

$$\boldsymbol{F}_{12} = \left(GM_1M_2/R^3\right)\boldsymbol{R}. \tag{19}$$

From Equations (18) and (19), an explicit expression for the parameter $\sigma'$ can be derived and also, with the help of Equations (11) and (12), the small mass to be attributed to the gravitational interaction can be found.

## 5. Discussion

Without claiming anything about the validity or the consequences of the model proposed, a simple straightforward model on the mutual interaction between two particles that influence the surrounding field leads elegantly to the right description of the gravitational interaction between two masses.

First, the Schrödinger equation is set up for two particles where the only assumption is that they are there and that the wave equation will have purely spherical symmetry. Surprisingly this is a solution that leads to the forming of pairs, in which only two particles are participating. One particular particle can form pairs with all others present, but each pair has to consist of two individual particles. The pair density probability, given by the wave function, is then used to substitute for the relativistically invariant entity, $m_0^2$, in the Dirac equation [1] [6]. It has been shown that the Dirac equation can be taken in the first calculation as well so that the whole analysis would be based on co-variant equations. This would lead to the same form of the probability density, although with different constants, but it creates no problems when allocating the mass dependences to the $\alpha$'s.

Replacing the quantities in the Dirac equation by the appropriate operators, and again taking purely spherical symmetry, it is found that a solution is only possible when the interaction energy of the pair is proportional to the inverse of the distance between them. It is important consider that the operators $\widehat{E_k}$ and $r_l$ commute. It means that "$Er$" is the quantity that has a sharp value, meaning that $E$ has sharp value if $r$ is well defined.

After adding up the effects the appropriate form of the classical Newton gravitation law is found.

The gravitational constant in Newton's law, $G$, is expressed by $G = \sqrt{6\sigma' c\hbar}$ in which the parameter $\sigma'$, equal to 2.34 × $10^4$ m/(kg)$^4$, can be seen as a universal constant that connects relativity with quantum mechanics.

The surprising, and at the same time bizarre, conclusion of the analysis given is that, apparently, each single particle has interaction with all other particles. It means that in the universe an unimaginable number of pair-wise interactions exists with greatly varying intensity and extensions and which depend on the masses of the members of the pair. It is difficult to comprehend, but it follows unambiguously from the equations describing the behaviour of the pairs.

An important aspect to mention is the fact that the right hand side in Equation (12) should be invariant under Lorentz transformation. However, the $r_{kl}$ is not. Therefore the parameters $\alpha_{kl}$ should transform in the same way as $r_{kl}$, but apparently it would make left and right hand side in equation 16 transform differently, which cannot be the case. We should however notice that the Planck's constant, $h$, is invariant, but $\hbar = h/2\pi$ is not.

Make the following "thought experiment". Consider a pair flying away from us at a speed $v$ such that the separation vector of the members of the pair is aligned in the direction of $v$. Due to the fact that π transforms just like $1/r_{kl}$ the result is that the interaction energy of the pair we measure becomes invariant. There is invariance throughout if the alignment is perpendicular to the speed. So the conclusion is that the interaction energy in the pair is invariant and independent of the alignment towards the observer, as it has to be.

Two remarks have to be made about the analysis proposed:

1) Equation (16) allows for both minus and plus signs for the E-values. It means that the force between particles can be negative and positive: repulsive and/or attractive. Apparently nature as we observe it has chosen for the low energy attractive variant. If the opposite would have been taken the universe could not exist.

2) The Equation (3) has singularities for $r_{kl}$ to zero. However, one might take the $\alpha_{ij}/r_0$ constant below a certain distance ($r_0$) from the particle centre and solve the Equation (8) for just one single particle and see that, differing from the analysis for pairs, a first order solution emerges and mass ($\hbar\alpha_{ij}/r_0 c$) is attributed to the particle, which to some extent is in analogy with superconductivity to explain the Meissner effect [1] [9] [10].

## 6. Conclusions

1) Using basic rules of quantum mechanics and relativity and preserving side conditions of invariance and co-variance where necessary, Newton's law on gravity can be derived theoretically. It contains the main factors of mass dependence and distance between particles and bodies in the right way.

2) It has by no means relevance to dark matter or dark energy. But the result allows for both attractive and/or repulsive gravitational interaction. It can be speculated that at certain distance or strength of the interaction it may change sign.

## Acknowledgements

The author thanks the Eindhoven University for giving the support to perform and publish this study.

Auke Veringa (averinga@versatel.nl) has been of great help to increase the accessibility of this paper also for non specialists in this field.

## References

[1] Ryder, L. (2009) Introduction to General Relativity. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9780511809033

[2] Heacox, W.D. (2015) The Expanding Universe. Cambridge University Press, Cambridge.

[3] Heitler, W. (1945) Elementary Wave Mechanics. Oxford Clarendon Press, Oxford.

[4] Ney, E.P. (1965) Electromagnetism and Relativity. Harper &Row, New York.

[5] Messiah, A. (1961) Quantum Mechanics. Vol. 1, North Holland Publishing Company, Amsterdam.

[6] Messiah, A. (1985) Quanten-Mechanik, Band 2. Walter de Gruyter, Berlin.

[7] Greulich, K.O. (2010) *Journal of Modern Physics*, **1**, 300-302. http://dx.doi.org/10.4236/jmp.2010.15042

[8] Panowsky, W. (1972) Melba Philips: Classical Electricity and Magnetism. Addison Wesley Publishing Company, Boston.

[9] Bardeen, J., Cooper, L.N. and Schrieffer, J.R. (1957) *Physical Review*, **108**, 1175. http://dx.doi.org/10.1103/PhysRev.108.1175

[10] Higgs, P.W. (1964) *Physics Review Letters*, **12**, 132-133. http://dx.doi.org/10.1016/0031-9163(64)91136-9

Journal of Modern Physics, 2016, 7, 1829-1844
http://www.scirp.org/journal/jmp
ISSN Online: 2153-120X
ISSN Print: 2153-1196

# A Possible Alternative to the Accelerating Universe III

**Frank R. Tangherlini**

San Diego, CA, USA
Email: frtan96@gmail.com

**How to cite this paper:** Tangherlini, F.R. (2016) A Possible Alternative to the Accelerating Universe III. *Journal of Modern Physics*, **7**, 1829-1844.
http://dx.doi.org/10.4236/jmp.2016.713163

**Received:** August 23, 2016
**Accepted:** September 27, 2016
**Published:** September 30, 2016



## Abstract

This work extends the author's two previous works (2015), *Journal of Modern Physics*, **6**, 78-87, and 1360-1370, by obtaining the index of refraction $n$ of the dark energy for additional values of the cosmological density parameters, and for the two methods of obtaining $n$: least squares fit, and electromagnetic theory. Comparison of the alternative model with the accelerating universe for the new values of the density parameters and $n$ is given in two tables. The new values for $n$ are used to obtain a range of ages for the Einstein de Sitter (EdS) universe. It is shown that the EdS universe must be older than the comparison accelerating universe. This requirement is met for the Planck 2015 value of the Hubble constant, corrected for the speed of light reduction by $n$. A supporting measurement as well as a disagreeing measurement is also discussed. Possible support from a stellar age determination is also discussed. It is shown that the expression obtained earlier for the increased apparent magnitude of the SNe Ia provides as good a fit for a closed universe with $\Omega(tot) = 1.005$, as it does for the flat EdS universe. Comparison is presented in a third table. An upper bound on $\Omega_\Lambda$ is given for a closed universe that eventually collapses back on itself that is too small for the value needed for the accelerating universe.

## Keywords

Dark Energy, Speed of Light, Age of Universe, Closed Universe

## 1. Introduction

In two previous works by the author [1] [2], hereafter referred to as I and II, it was shown that it is possible to explain the diminished brightness of the Type Ia supernovae (SNe Ia) found by Perlmutter *et al.* [3] [4], Riess *et al.* [5], and Schmidt *et al.* [6], and the increased distance to the "standard ruler" of the baryon acoustic oscillations (BAO) determined by Anderson *et al.* [7] [8], by assuming that the speed of light through the dark energy of intergalactic space has been reduced to *c/n*, where *n* is the index of refraction of the dark energy. Thus, in this alternative model, the dark energy no longer has associated with its energy-momentum source tensor a negative pressure that causes the expansion of the universe to accelerate (for a review see, e.g., Wang [9]), but instead has an index of refraction greater than unity. It was also assumed that the dark energy is another phase of dark matter, and that the phase transformation started to take place at about redshift $z = 1.65 \pm 0.15$, as discussed in Riess *et al.* [10], where there appears to be a supernova not exhibiting acceleration, and in the proposed alternative model, is where the dark energy started to appear as a consequence of the expansion cooling of the dark matter that was present in intergalactic space. Since that expansion cooling did not take place for the dark matter associated with the galaxies, because it is the space between the galaxies that expands, not the galaxies themselves, hence within the galaxies, *n* remains unity, and the speed of light is *c*. Since galaxies do not have a sharp boundary, with the dark matter halos extending well beyond the central luminous baryonic regions, there will be a transition region where *n* changes from unity to its intergalactic value, but for simplicity this is ignored at this stage of the study.

The purpose of this work is to extend the previous work in I and II by making use of additional values of the cosmological density parameters to obtain new values for *n*, and to use these new values to obtain additional estimates of the age of the universe according to the alternative model.

Also, since the comparisons in I and II have been between the ΛCDM accelerating universe, and only the flat Einstein de Sitter universe, for completeness, an additional comparison is made here with a closed universe.

In Section 2, the new values of *n* will be obtained by using both the least squares method that was introduced in I, and the square root method based on electromagnetic theory that was introduced in II. Both methods will be based on the values of the cosmological density parameters from the Planck satellite work of Abe *et al.* [11]. A comparison with the prediction for the increased distance to the SNe Ia, and the BAO standard ruler for different values of the cosmological density parameters is presented in **Table 1** and **Table 2**. In Section 3, the range of values for *n* obtained in Section 2 is used to obtain a range of ages for the universe. All of these ages are greater than or equal to 14 Gyr, and hence exceed that for the accelerating universe of 13.8 ± 0.1 Gyr as given in [11]. Significantly, a qualitative argument is presented that shows that the age of the comparison Einstein de Sitter universe has to be greater than that of the accelerating universe, as was already found quantitatively in II. Possible empirical supports, as well as a possible empirical objection to the alternative model, are also discussed. In Section 4, as mentioned above, instead of comparing the accelerating universe with the flat Einstein de Sitter universe, the comparison is made with a closed universe. In I it was briefly noted that the expression that was found there for the increase in the apparent magnitude of the SNe Ia did not require the universe to be flat, and in this section, this is demonstrated for a closed university with a total density parameter,

$\Omega = 1.005$, which is the one sigma upper bound given in [11]. It is also shown that fo- this case, both the least squares value of $n$, and the age of the closed universe, are the same (to three places) as that for the Einstein-de Sitter universe. In Table 3, a comparison is made between the closed universe and the flat universe for their percentage fit with the accelerating universe. An upper bound on $\Omega_\Lambda$ for a closed universe that eventually collapses back on itself is also derived and discussed. In Section 5, there are concluding remarks.

## 2. Additional Determinations of *n*

In I, where the cosmological density parameters were given by $\Omega_m = 0.30$, $\Omega_\Lambda = 0.70$, the least squares value of $n$ was found to be $n = 1.49$. In II, upon setting $\Omega_\Lambda = \Omega_{de}$ (where the subscript "de" refers to dark energy which in the alternative model has the same energy density as the cosmological term, but, in contrast, has negligible stress), and using the electromagnetic relation, $n = (KK_\mu)^{1/2}$, where $K$ is the dielectric constant, and $K_\mu$ is the relative permeability of the dark energy, and the assumption that $KK_\mu = \Omega_{de}/\Omega_m$ so that one has another method of determining $n$ given by $n = (\Omega_{de}/\Omega_m)^{1/2}$, it was found for the above density parameters that $n = 1.53$. In what follows, it will be convenient to denote the least squares value for $n$ as $n(ls)$, and to denote the value for $n$ obtained from the square root as $n(sr)$. It was also found in II that $n(ls) = 1.47$ and $n(sr) = 1.46$, upon employing the one sigma upper limit of $\Omega_m = 0.308 \pm 0.012$ given in [11], so that $\Omega_m = 0.32$, and $\Omega_\Lambda = 0.68$. In this section, values of $n(ls)$ and $n(sr)$ will be obtained for the mean value $\Omega_m = 0.308$, and the one sigma lower limit $\Omega_m = 0.296$, with $\Omega_m + \Omega_\Lambda = 1$ as before, so that $\Omega_\Lambda = 0.692$, and $\Omega_\Lambda = 0.704$, respectively.

For the reader's convenience, the analysis leading to $n(ls)$ will next be briefly reviewed. As derived in Section 6 of I, and briefly recapitulated in II, $\mathrm{X}_\Lambda(z)$ is proportional to the distance in the accelerating universe out to an object at redshift $z$, while $\mathrm{X}_m(z)$ is proportional to the distance for the Einstein de Sitter universe. (See also the discussion in Section 4 below, where more details are given.) The general expression for $\mathrm{X}(z)$ is

$$\mathrm{X}(z) = \int_0^z \mathrm{d}z'/E(z'), \tag{1}$$

with

$$E(z) \equiv H(z)/H_0 = (\Omega_m)^{1/2}\sqrt{(1+z)^3 + (\Omega_\Lambda/\Omega_m)}, \tag{2}$$

where $H(z) \equiv \dot{a}/a$ is the Hubble parameter, and $H_0$, its value at the present epoch, the Hubble constant, and where $a$ is the FLRW expansion parameter, and also $a(z) = a_0/(1+z)$. From (1) and (2), as given in I and II, it follows that

$$\mathrm{X}_\Lambda(z) = (\Omega_m)^{-1/2}\int_0^z \mathrm{d}z'\Big/\sqrt{(1+z)^3 + (\Omega_\Lambda/\Omega_m)}, \tag{3}$$

so that for the values $\Omega_m = 0.308, \Omega_\Lambda = 0.692$, one has

$$\mathrm{X}_\Lambda = (0.308)^{-1/2}\int_0^z \mathrm{d}z'\Big/\sqrt{(1+z')^3 + 2.247}. \tag{4}$$

As noted in I and II, this integral has to be evaluated numerically. In contrast, for the Einstein de Sitter universe, with $\Omega_m = 1, \Omega_\Lambda = 0$, one has the same value for $X_m(z)$ as was found in I and II, for which the integral is immediate, and given by

$$X_m(z) = \int_0^z dz'(1+z')^{-3/2} = 2\left(1-(1+z)^{-1/2}\right). \tag{5}$$

As in I and II, the difference of the logarithmic proportional distances between the accelerating universe and the Einstein de Sitter universe is given by

$$\log X_\Lambda(z) - \log X_m(z) = \log\left(X_\Lambda(z)/X_m(z)\right) = d, \tag{6}$$

where

$$d = \log\left(1+(n-1)\ln(1+z)\right) \tag{7}$$

is the prediction of the alternative model under the assumption (for which there is as yet no theoretical foundation) that the light speed has been reduced to $c/n$ when traveling through the dark energy of intergalactic space, as shown in I, Section 4, page 82, and where the increase in apparent magnitude $\delta m$ is shown to be given by $\delta m = 5d$. For the percentage comparisons, it is sufficient to work with $d$ rather than $\delta m$. For $\Omega_m = 0.308$, $\Omega_\Lambda = 0.692$ it was found for the redshift range, $0 \le z \le 1.0$, that $n(ls) = 1.48$. Another value for n is obtained under the assumption that it is of electromagnetic origin, and as discussed in II, and above, so that $n(sr) = (KK_\mu)^{1/2} = (\Omega_{de}/\Omega_m)^{1/2}$. Hence, for $\Omega_m = 0.308, \Omega_{de} = \Omega_\Lambda = 0.692$, one has $n(sr) = 1.50$. In Table 1, a comparison is given for the percentage disagreement with the accelerating universe for $n(ls)$ and $n(sr) = 1.50.$, this is analogous to Table 1 in II, and for just $n(ls)$, to Table 3 in I.

Table 1. Comparison of $\log\left(X_\Lambda(z)/X_m(z)\right)$ with $d = \log\left(1+(n-1)\ln(1+z)\right)$ for $n(ls) = 1.48$ and $n(sr) = 1.50$, $\Delta \equiv d - \log\left(X_\Lambda(z)/X_m(z)\right)$, and for brevity, $R \equiv \log\left(X_\Lambda/X_m\right)$, and the arguments for $R$, $X_\Lambda, X_m$ are omitted. $\Omega_m = 0.308, \Omega_\Lambda = 0.692$, $\Omega_\Lambda/\Omega_m = 2.247$.

| $z$ | $\log(X_\Lambda/X_m)$ | $d(1.48)$ | $\Delta(1.48)$ | $\Delta(1.48)/R\%$ | $d(1.50)$ | $\Delta(1.50)$ | $\Delta(1.50)/R\%$ |
|---|---|---|---|---|---|---|---|
| 0.1 | 0.0208 | 0.0194 | −0.0014 | −6.7 | 0.0202 | −0.0006 | −2.9 |
| 0.2 | 0.0387 | 0.0364 | −0.0023 | −5.9 | 0.0379 | −0.0008 | −2.1 |
| 0.3 | 0.0539 | 0.0515 | −0.0024 | −4.5 | 0.0535 | −0.0004 | −0.7 |
| 0.4 | 0.0671 | 0.0650 | −0.0021 | −3.1 | 0.0675 | 0.0004 | 0.6 |
| 0.5 | 0.0786 | 0.0772 | −0.0014 | −1.8 | 0.0802 | 0.0016 | 2.0 |
| 0.6 | 0.0885 | 0.0883 | −0.0002 | −0.2 | 0.0917 | 0.0032 | 3.6 |
| 0.7 | 0.0972 | 0.0985 | 0.0013 | 1.3 | 0.1022 | 0.0050 | 5.1 |
| 0.8 | 0.1048 | 0.1079 | 0.0031 | 3.0 | 0.1119 | 0.0071 | 6.8 |
| 0.9 | 0.1115 | 0.1166 | 0.0051 | 4.6 | 0.1209 | 0.0094 | 8.4 |
| 1.0 | 0.1175 | 0.1247 | 0.0072 | 6.1 | 0.1292 | 0.0117 | 10.0 |

In Table 2, the percentage disagreement results are given for $\Omega_m = 0.296$, and $\Omega_\Lambda = 0.704$. It was found that for this case, $n(ls) = 1.50$ and $n(sr) = 1.54$. Hence with Table 1 and Table 2 here, and Table 1 in II, the totality of cases are presented for $\Omega_m = 0.308 \pm 0.012$, along with the corresponding values for $\Omega_\Lambda$.

Table 2. Comparison of $\log\left(X_\Lambda(z)/X_m(z)\right)$ with $d = \log\left(1+(n-1)\ln(1+z)\right)$ for $n(ls) = 1.50$ and $n(sr) = 1.54$, $\Delta \equiv d - \log\left(X_\Lambda(z)/X_m(z)\right)$, and for brevity, $R \equiv \log\left(X_\Lambda/X_m\right)$, , and the arguments for $R, X_\Lambda, X_m$ are omitted. $\Omega_m = 0.296, \Omega_\Lambda = 0.704$, $\Omega_\Lambda/\Omega_m = 2.378$.

| $z$ | $\log(X_\Lambda/X_m)$ | $d(1.50)$ | $\Delta(1.50)$ | $\Delta(1.50)/R\%$ | $d(1.54)$ | $\Delta(1.54)$ | $\Delta(1.54)/R\%$ |
|---|---|---|---|---|---|---|---|
| 0.1 | 0.0210 | 0.0202 | −0.0008 | −3.8 | 0.0218 | 0.0008 | 3.8 |
| 0.2 | 0.0394 | 0.0379 | −0.0015 | −3.8 | 0.0408 | 0.0014 | 3.6 |
| 0.3 | 0.0551 | 0.0535 | −0.0016 | −2.9 | 0.0575 | 0.0024 | 4.4 |
| 0.4 | 0.0686 | 0.0675 | −0.0011 | −1.6 | 0.0725 | 0.0039 | 5.7 |
| 0.5 | 0.0804 | 0.0802 | −0.0002 | −0.2 | 0.0860 | 0.0056 | 7.0 |
| 0.6 | 0.0906 | 0.0917 | 0.0011 | 1.2 | 0.0982 | 0.0076 | 8.4 |
| 0.7 | 0.0996 | 0.1022 | 0.0026 | 2.6 | 0.1094 | 0.0098 | 9.8 |
| 0.8 | 0.1074 | 0.1119 | 0.0045 | 4.2 | 0.1197 | 0.0123 | 11.5 |
| 0.9 | 0.1144 | 0.1209 | 0.0065 | 5.7 | 0.1292 | 0.0148 | 12.9 |
| 1.0 | 0.1206 | 0.1292 | 0.0086 | 7.1 | 0.1381 | 0.0175 | 14.5 |

It is clear from the last column in Table 2, in which, particularly at $z = 0.5$, the disagreement is 7.0%, that $n(sr) = 1.54$ is ruled out. However, this need not mean that the square root method of obtaining $n$ is at fault, rather it could be that the problem is with the values of the density parameters, *i.e.*, $\Omega_m = 0.296, \Omega_\Lambda = 0704$. If one assumes that $n(sr) = \sqrt{\Omega_{de}/\Omega_m}$ yields a valid determination of $n$ for, say, redshifts, $0 \le z \le 0.7$, then it follows that the large percentage disagreement for $n(sr) = 1.54$ is indicative that the above values of the cosmological density parameters are ruled out instead. If one demands that the percentage disagreement not exceed |2%|, for $z = 0.5$, then from Table 1, $\Omega_m = 0.308$ provides a lower bound for $\Omega_m$, while for an upper bound, it was found that for $\Omega_m = 0.317$ the disagreement at $z = 0.5$ is −1.8%, while for $\Omega_m \ge 0.318$, the disagreement is less than −2%, it follows that $\Omega_m = 0.317$ provides an upper bound. Hence, allowing $\pm 2\%$ disagreement, the requirement that the square root method is valid, can be seen as predicting

$$0.308 \le \Omega_m \le 0.317. \tag{8}$$

Also, since for $\Omega_m = 0.317$, $\Omega_{de} = \Omega_\Lambda = 0.693$, one has that $n(sr) = 1.47$. The corresponding range of values of $n$, that include least squares values as well as square root values is given by $n = 1.48^{+0.02}_{-0.01}$. However, in calculating $n(ls)$, for $\Omega_m = 0.308$, if one restricts the redshifts to $z \le 0.7$, one finds that $n(ls) = 1.49$. Thus a possibly better estimate of $n$ is given by

$$n = 1.49^{+0.01}_{-0.02}. \tag{9}$$

Indeed, for $n$ = 1.49, at $z$ = 0.5, one finds from **Table 1** that $\Delta = 0.0001$, and hence $\Delta/R = 0.1\%$.

It will be noticed from the tables that lower values of $n$ fit better for higher values of $z$, while higher values of $n$ fit better for lower values of $z$. Interestingly, the disagreement for the lower values of $z$, *i.e.* $z \le 0.3$, suggest that if the alternative, model is correct, it predicts a slightly brighter SNe Ia than the accelerating ΛCDM universe for this lower range of redshifts. As noted in I, this could provide another means for choosing between the two theories. On the other hand, the percentage disagreements for, say, $z > 0.7$, are probably an indication that the simplifying assumption that the index of refraction is constant for these higher values of $z$ is not valid, and that at these redshifts, and higher, the dark matter has not fully transformed into dark energy. Thus a future, more accurate model should assume that $n = n(z)$, with the requirement that $n(z)$ approach unity for sufficiently high $z$. As noted above, at $z = 1.65 \pm 0.15$ [10], the universe no longer seems to be accelerating, and in the present model, this should be the approximate redshift where the intergalactic dark matter began to make a phase transition into dark energy, and the index of refraction started to increase from unity. But one will need to know more about the properties of dark matter and dark energy to go further. Indeed, it will be noticed that in this work, no assumptions have been made about the particle nature of the dark matter, nor the dynamical nature of the phase transition into dark energy. At this stage of the investigation, the model is purely phenomenological, and it may be possible to determine $n(z)$ phenomenologically, which could then be helpful in probing the nature of the dark matter, and the proposed phase transition into dark energy.

## 3. Age of the Universe

In II it was found that for $n = 1.46$, the age of the Einstein de Sitter universe, with $H_0 = 67.8 \pm 0.9\ \text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$ [11], which was rounded to 68 km·s$^{-1}$·Mpc$^{-1}$, was 14.0 Gyr, which was obtained in II, Section 5, (17), and is equivalent to (11) below. In this section, a range of ages for the Einstein de Sitter universe based on the range of values of $n$ given in Section 2 will be presented here, and discussed. The above age of 14.0 Gyr, and those that will be obtained below, are obviously greater than the age of the accelerating universe of 13.8 ± 0.1 Gyr, given in [11], and hence it is of interest to present a qualitative argument to show why the age of the Einstein de Sitter universe in the alternative model is necessarily greater than the age of the accelerating universe. To accomplish this, it is convenient to divide the expansion of the universe into three eras, based on the different speed of light in these three eras, as discussed below: the first era is the expansion from the big bang to the time when the universe started to accelerate. At that time, the expansion parameter would be somewhat larger for the accelerating universe than that for the Einstein de Sitter universe, since while both universes had been decelerating, the former's deceleration would have diminished more rapidly to zero. Call this value of the expansion parameter for the accelerating universe $a'$. Therefore the time for the Einstein de Sitter universe to expand to $a'$ would be greater than that for the accelerating universe, call this time difference $\Delta t_1$, with $\Delta t_1 > 0$. The second era is the interval of time in which the universe expanded from $a'$ to its size at redshift $z = 0.5$, call this value of the expansion parameter $a''$. It would require a longer amount of time for the Einstein de Sitter universe to expand from $a'$ to $a''$ than for the accelerating universe, so that the former would require an additional time $\Delta t_2$ with $\Delta t_2 > 0$. The third era would be the expansion from $z = 0.5$ to the present epoch $z = 0$, and again the accelerating universe would reach the present value of the expansion parameter $a_0$ in a shorter time than the decelerating universe which would require an additional time $\Delta t_3 > 0$. During this last era, the speed of light according to the alternative model is $c/n$, and hence the increased length of time the light takes to travel from the SNe Ia to earth, over what it would be if it traveled at speed $c$, leads to the increase in time needed for the decelerating Einstein de Sitter universe to expand the extra distance needed to explain the increase in apparent magnitude of the SNe Ia at $z = 0.5$, over what it would have been if light had traveled with speed $c$. On the other hand, during the second epoch, the speed of light through the dark energy of intergalactic space is gradually changing from $c$ to $c/n$; however, the details of this change, together with the postulated phase change of dark matter into the dark energy of intergalactic space is left to future studies. In any case, it is clear that the sum $\Delta t_1 + \Delta t_2 + \Delta t_3$ leads to a greater total time back to the big bang for the Einstein de Sitter universe than for the accelerating universe, as was already found in the particular case considered above, for $n$ 1.46; but, as demonstrated above, the result is quite general.

Now, as was discussed in II, because the determination of the Hubble constant involves the first order Doppler effect for the light that has traveled through intergalactic space, and since the distances involved are for $z < 0.5$ in which, according to the model, the speed of light is $c/n$, one has to correct the Doppler expression to allow for this, so that it becomes $\lambda_0 = \lambda\left(1 + \left(nv/c\right)\right)$, where $\lambda_0$ is the wavelength observed at the present epoch, $\lambda$ is the wavelength in the rest frame of the receding galaxy, and $v$ is the Hubble flow recession velocity of the galaxy. With the red shift $z \equiv \left(\lambda_0 - \lambda\right)/\lambda$, one has for $n = 1$, $cz = v$, and the basic discovery of Hubble is that $v = H_0 D$, where $D$ is the proper distance to the galaxy at $z$, so that the standard expression is $cz = H_0 D$, from which one determines the Hubble constant as $H_0 = cz/D$. But when one takes into account the reduced speed of light, the new corrected value of the Hubble constant, denoted by $H_0^*$, is given by

$$H_0^* = cz/nD = H_0/n. \tag{10}$$

Since the age of the Einstein de Sitter universe, taken to be from the big bang to the present epoch, and denoted by $T_0$, is given by $\left(2/3\right)\left(\dot{a}/a\right)_0^{-1}$, for $n = 1$, this becomes $T_0 = 2H_0^{-1}/3$. However, as remarked in II, for $n \neq 1$, as is the case here, this has to be corrected to an age $T_0^*$ given by

$$T_0^* = \frac{2}{3}\left(H_0^*\right)^{-1} = \frac{2}{3} n H_0^{-1}. \tag{11}$$

Since (11) is indifferent to whether one works with $n(ls)$, or $n(sr)$, upon introducing the value of $n$ in (9), and with $H_0^{-1} = 14.4 \pm 0.2\ \text{Gyr}$ based on $H_0$ = 67.8 ± 0.9 km·s$^{-1}$·Mpc$^{-1}$ [11] one has that

$$T_0^* = 14.3 \pm 0.3\ \text{Gyr}. \tag{12}$$

This clearly exceeds the age of the accelerating universe of $13.8 \pm 0.1\ \text{Gyr}$, as was expected on the basis of the qualitative analysis given above. In this case, $\Delta t_1 + \Delta t_2 + \Delta t_3$ = 0.5 ± 0.3 Gyr.

With regard to this revision of the Hubble constant, it should be pointed out that since the light from, say, the Cepheid variables that is used in determining $H_0$, passes through the host galaxy as well as our own Galaxy, where in both cases the speed of light is $c$, this has the consequence that the effective speed of light for the entire path is greater than $c/n$. However a rough estimate indicates that a correction for this effect would be less than 0.5 percent, and in view of the sizes of the other uncertainties, it can be ignored at this stage of the study.

An interesting hint of possible support for the alternative model comes from the fairly recent study of the star HD 140283 by Bond *et al.* [12], who describe it as a subgiant with low metallicity in the solar neighborhood, at approximately 100 lyr from earth. They found that when all uncertainties are included the star's age is 14.46 ± 0.8 Gyr. They emphasized that this age was not in disagreement with the age of the universe when allowance was made for the uncertainty in the star's age. At that time, 2013, no other age than that of the accelerating universe (which was then given as 13.77 ± 0.06 Gyr) was available for comparison, but clearly the mean age they found for the star puts its age closer to $T_0^*$ than to that of the accelerating universe. However, until the substantial uncertainty in the star's age has been reduced, obviously no firm conclusion can be drawn. But it is noteworthy that, in addition to the divergent lensing possibility described in Section 4 of II, the difference in the ages of the universe for the two models is yet another avenue of approach for deciding between the two models.

Quite recently there has appeared the latest finding of Riess *et al.* [13] that yields a 2.4% determination of the local value of the Hubble constant of $73.24 \pm 1.74\ \text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$. This leads to a Hubble time to three places given by $H_0^{-1} = 13.4 \pm 0.3\ \text{Gyr}$. Hence, according to the alternative model, since the age of the universe, as given in (11) is $(2/3) n H_0^{-1}$, upon introducing the largest value of $n$ obtained from (9), *i.e.* $n = 1.50$, one obtains an age $T_0^*$ for the Einstein de Sitter universe of $13.4 \pm 0.3\ \text{Gyr}$. Since, as was pointed out earlier in this section, the age of the Einstein de Sitter universe has to be greater than that of the accelerating universe, hence greater than $13.8 \pm 0.1\ \text{Gyr}$, their value for $H_0$ is possibly in conflict with the alternative model. On the other hand, the possible conflict is less than with the 3.3% determination by Riess *et al.* [14] of 2011,for which $H_0 = 73.8 \pm 2.4\ \text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$. Interestingly, Cheng and Huang [15], on the basis of their BAO studies, found that $H_0 = 68.11 \pm 0.86\ \text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$ which is in excellent agreement with the Planck [11] finding that $H_0 = 67.8 \pm 0.9\ \text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$, the value used here. They discussed their disagreement with the 2011 value of $H_0$ found by Riess *et al.* [14]. Since the alternative model also disagrees with this value given in [14], and to a lesser degree with their more recent value [13], this could be a sign that the alternative model is possibly on the right track, and has predictive powers. However, as noted above, the large uncertainties surrounding the various values make it impossible to draw scientifically valid conclusions as to agreement or disagreement.

It has undoubtedly been noticed that the value of $n$ is very nearly—if not exactly—the reciprocal of the two-thirds factor relating the age of the Einstein de Sitter universe to the Hubble time. At this writing, there is no explanation for this unexpected relation, and it is left to future studies to find a possible solution.

## 4. Comparison with a Closed Universe

As mentioned in I, the expression that was derived therein Section 4, p.82, for the increased apparent magnitude of the SNe Ia given by

$$\delta m = 5\log\left(1+\left(n-1\right)\ln\left(1+z\right)\right), \tag{13}$$

does not just hold for the flat Einstein de Sitter universe, it is applicable to any isotropic universe for which the expansion parameter satisfies $a = a_0/\left(1+z\right)$, which is a generic relationship for FLRW expanding universes. In view of the long standing interest in closed universes, and the author's work on a closed pulsating universe [16] [17], it is appropriate to determine what values of $n$ emerge from comparison with a closed universe. As pointed out in [11], the equivalent cosmological density parameter for curvature differs from zero by an amount given by $\Omega_k = \pm 0.005$. However, it should be emphasized that this value represents a measurement uncertainty, and the true value of the curvature parameter could be much smaller, and even zero, as predicted by the inflationary models of Guth [18] [19], Linde [20], and Albrecht and Steinhardt [21]. This possible curvature contribution to $\Omega$ in the author's approach is written differently, and will now be derived.

The standard line element for a general, homogeneous, isotropic, time-orthogonal, FLRW universe, in isotropic coordinates, can be written as

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a\left(t\right)^2\left(1+\left(kr^2/4\right)\right)^{-2}\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j,\ \left(i, j = 1,2,3\right) \tag{14}$$

with $r^2 = \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j$, and $k = 1,0,-1$ for a closed, flat, and open universe, respectively. The Einstein field equation for the energy density $G_0^0 = -\kappa T_0^0$, in the absence of CMB radiation, neutrino background, and the cosmological term, reduces to

$$\dot{a}^2 - \left(8\pi G\rho a^2/3\right) = -kc^2. \tag{15}$$

For the Einstein de Sitter universe $k = 0$, while in this section $k = 1$. Hence this equation can be rewritten as

$$1+\frac{c^2}{\dot{a}^2} = \frac{8\pi G\rho a^2}{3\dot{a}^2}. \tag{16}$$

Since $\rho_0 a_0^3 = \rho a^3$ from the covariant conservation law, and since $a = a_0/(1+z)$, the above becomes

$$1+\frac{c^2}{\dot{a}^2}=\frac{8\pi G\rho_0(1+z)^3 a^2}{3\dot{a}^2}. \tag{17}$$

Hence, at the present epoch, $z = 0$, with $a = a_0, \dot{a} = \dot{a}_0$, and $H_0 \equiv \dot{a}_0/a_0$, one has

$$1+\frac{c^2}{\dot{a}_0^2}=\frac{8\pi G\rho_0}{3H_0^2}=\frac{\rho_0}{\rho_c}, \tag{18}$$

where $\rho_c \equiv 3H_0^2/8\pi G$. And since $\Omega_m \equiv \rho_0/\rho_c$, in contrast with the Einstein de Sitter universe, for which $\Omega_m = 1$, one has instead

$$\Omega_m = 1+\frac{c^2}{\dot{a}_0^2}. \tag{19}$$

One now proceeds as in I and II. From (17), and the relations leading to (19), one has

$$\frac{H}{H_0}=\frac{\dot{a}}{H_0 a}=\sqrt{\frac{8\pi G\rho_0(1+z)^3}{3H_0^2}-\frac{c^2(1+z)^2}{a_0^2 H_0^2}}. \tag{20}$$

Then from $\Omega_m \equiv \rho_0/\rho_c = 8\pi G\rho_0/3H_0^2$, and the definition $E(z) \equiv H(z)/H_0$, upon simplification, (20), becomes

$$E(z)=(\Omega_m)^{1/2}(1+z)\sqrt{(1+z)-\left(c^2/\dot{a}_0^2\Omega_m\right)}. \tag{21}$$

Upon replacing $\Omega_m$ by $\left(1+\left(c^2/\dot{a}_0^2\right)\right)$ from (19), the above may be rewritten as

$$E(z)=\left(1+\left(c^2/\dot{a}_0^2\right)\right)^{1/2}(1+z)\sqrt{(1+z)-\left(c^2/\dot{a}_0^2\left(1+\left(c^2/\dot{a}_0^2\right)\right)\right)}. \tag{22}$$

The proportional distance function X($z$), as previously mentioned, is defined as

$$\mathrm{X}(z) \equiv \int_0^z \mathrm{d}z'/E(z'). \tag{23}$$

Now $\mathrm{X}_\Lambda(z)$ will be the same as previously used in comparison with the flat universe, here, on the other hand, it will be compared with the new form for $\mathrm{X}_m(z)$ which, from (22), is given by

$$\mathrm{X}_m(z)=\left(1+\left(c^2/\dot{a}_0^2\right)\right)^{-1/2}\int_0^z\frac{\mathrm{d}z'}{(1+z')\sqrt{(1+z')-\left(c^2/\dot{a}_0^2\left(1+\left(c^2/\dot{a}_0^2\right)\right)\right)}}. \tag{24}$$

Upon denoting the ratio $c^2/\dot{a}_0^2$ by $f$, the above expression can be written as

$$\mathrm{X}_m(z)=(1+f)^{-1/2}\int_0^z\frac{\mathrm{d}z'}{(1+z')\sqrt{(1+z)-f(1+f)^{-1}}}. \tag{25}$$

After combining the factor $2f^{-1/2}(1+f)^{1/2}$ that emerges from the integral with the pre-factor in (25), the above integration yields

$$\mathrm{X}_m(z)=2f^{-1/2}\left\{\tan^{-1}\sqrt{\frac{(1+z)-f(1+f)^{-1}}{f(1+f)^{-1}}}-\tan^{-1}f^{-1/2}\right\}. \tag{26}$$

The following cosmological density parameters, $\Omega_m = 0.308, \Omega_\Lambda = 0.692$ were chosen for the accelerating universe to obtain $X_\Lambda$ for comparison with the closed universe. Thus, the results presented in Table 3 below will be similar to that of Table 1, except $X_m$ will now be that for a closed universe. However, in Table 1, in comparison with the flat universe, $n(ls) = 1.48$ was used, and so, before making the comparison, it is necessary to determine $n(ls)$ for the closed universe. For $f \equiv c^2/\dot{a}_0^2 = 0.005$, it was found $n(ls)$ had the same value as for the flat universe, *i.e.*, 1.48. In Table 3, the last column is taken from the fifth column in Table 1.

Table 3. Comparison of $\log(X_\Lambda(z)/X_m(z))$, where $X_m(z)$ is for a closed universe, with the logarithmic distance correction $d = \log(1+(n-1)\ln(1+z))$ for $n = 1.48$, $\Delta \equiv d - \log(X_\Lambda/X_m)$. For brevity, $R \equiv \log(X_\Lambda/X_m)$, and the arguments for $\Delta$, $R$, $X_\Lambda$, $X_m$ are omitted.

| $z$ | $\log(X_\Lambda/X_m)$ | $d$ | $\Delta$ | $\Delta/R\%$ | $\Delta/R\%(flat)$ |
|---|---|---|---|---|---|
| 0.1 | 0.0208 | 0.0194 | −0.0014 | −6.7 | −6.7 |
| 0.2 | 0.0387 | 0.0364 | −0.0023 | −5.9 | −5.9 |
| 0.3 | 0.0540 | 0.0515 | −0.0025 | −4.6 | −4.5 |
| 0.4 | 0.0673 | 0.0650 | −0.0023 | −3.4 | −3.1 |
| 0.5 | 0.0787 | 0.0772 | −0.0015 | −1.9 | −1.8 |
| 0.6 | 0.0887 | 0.0883 | −0.0004 | −0.5 | −0.2 |
| 0.7 | 0.0974 | 0.0985 | 0.0011 | 1.1 | 1.3 |
| 0.8 | 0.1050 | 0.1079 | 0.0029 | 2.8 | 3.0 |
| 0.9 | 0.1118 | 0.1166 | 0.0048 | 4.3 | 4.6 |
| 1.0 | 0.1178 | 0.1247 | 0.0069 | 5.9 | 6.1 |

As can be seen from the table, the fit for the closed universe is very nearly the same as for the flat Einstein de Sitter universe. Moreover, as will be shown next, the age of the closed universe is very nearly the same as for the flat universe for this value of $\Omega_m$ that is so close to unity.

It is convenient to rewrite (15) as

$$\frac{\dot{a}^2}{2} - \frac{GM}{a} = -\frac{c^2}{2}, \tag{27}$$

where $M \equiv (4\pi/3)\rho a^3$, although, to be sure, for a closed, spherical universe, the actual mass is actually $2\pi^2 \rho a^3$. However, in the following, one never has to make explicit use of the mass, and because of the obvious similarity to Newtonian mechanics, (27) is simpler to work with, since the solution to the differential equation is well known, and is given by the parametric equations for a cycloid

$$a = (GM/c^2)(1 - \cos\varphi), \tag{28}$$

$$t = \left(GM/c^3\right)\left(\varphi - \sin\varphi\right). \tag{29}$$

From these two equations one obtains an expression for $\dot{a}$ given by

$$\dot{a} = c\sin\varphi/\left(1-\cos\varphi\right), \tag{30}$$

from which, together with (28), one obtains the Hubble parameter $H \equiv \dot{a}/a$ in the form

$$H = \frac{c^3\sin\varphi}{GM\left(1-\cos\varphi\right)^2}. \tag{31}$$

For $\varphi \ll 1$, from (29), $t = GM\varphi^3/6c^3$, and from (31), $H = 4c^3/GM\varphi^3$, hence

$$tH = 2/3, \tag{32}$$

which is the same relation that holds for the Einstein de Sitter universe. For the larger value of $\varphi$ that is needed here, one expands $t$ and $H$ to the next higher order terms. After combining (29) and (31), one has

$$tH = \frac{\sin\varphi\left(\varphi - \sin\varphi\right)}{\left(1-\cos\varphi\right)^2}. \tag{33}$$

After setting $\sin\varphi = \varphi - \varphi^3/6 + \varphi^5/120$, and $\varphi - \sin\varphi = \varphi^3/6 + \varphi^5/120$, upon neglecting higher order terms, and then taking the product, the numerator in (33) becomes $\left(\varphi^4/6\right)\left(1-\left(13\varphi^2/60\right)\right)$, after further neglect of higher order terms. After approximating the denominator as $\left(\varphi^4/4\right)\left(1-\left(\varphi^2/12\right)\right)$, and upon further neglect of higher order terms, one finds

$$tH = \frac{2}{3}\left(1-\frac{\varphi^2}{20}\right). \tag{34}$$

To determine $\varphi$ in terms of the departure $\varepsilon$ of $\Omega_m$ from its value for the flat Einstein de Sitter universe, one uses from (19) that $\Omega_m = 1 + c^2/\dot{a}^2$, so that $\varepsilon = c^2/\dot{a}^2$, and one will determine $\varphi$ in terms of $\varepsilon$. From (30), since $c^2/\dot{a}^2 = \left(1-\cos\varphi\right)^2/\sin^2\varphi$, one obtains

$$\Omega_m = \frac{2}{1+\cos\varphi}. \tag{35}$$

After substituting the approximation $\cos\varphi = 1 - \varphi^2/2 + \varphi^4/24$ in (35) and expanding upon further neglect of higher order terms, one obtains

$$\Omega_m = 1 + \varphi^2/4 + \varphi^4/24. \tag{36}$$

Upon setting $\Omega_m = 1 + \varepsilon$ in (36), there results a quadratic equation for $\varphi^2$ given by

$$\varphi^4 + 6\varphi^2 - 24\varepsilon = 0. \tag{37}$$

The positive root for $\varphi^2$ to first order in $\varepsilon$ is $\varphi^2 = 4\varepsilon$. Hence, for this level of approximation, (34) becomes

$$tH = \frac{2}{3}\left(1-\frac{\varepsilon}{5}\right). \tag{38}$$

With $\varepsilon = 0.005$, one has $t = 0.666 H^{-1}$, and hence for this case, as well as for closed universes of even lesser curvature, the age difference between a closed universe and the Einstein de Sitter universe is negligible. Some further observations concerning the two models are of interest.

Although for the closed universe, from (30), $\dot{a} > c$, for $\varphi < \pi/2$, this does not violate $c$ being the limiting speed, since $\dot{a}$ is analogous to the Minkowski speed in special relativity, which can be arbitrarily large, since the speed of light determined in this way is infinite. In other words, the coordinate time $t$ in the FRLW line element is analogous to a proper time; for further discussion, see that by the author in [22]. Also, it should be noted that the solution for $a(t)$ for the Einstein de Sitter universe, *i.e.*, $a(t) \propto t^{2/3}$, which is usually described as having the time run $0 \le t \le \infty$, also holds for $t < 0$, so one may think of the time running for the complete solution as $-\infty \le t \le \infty$. In the negative branch, for which the time runs $-\infty \le t < 0$, one has a universe collapsing from infinity to a cusp at $t = 0$, the big bang, where $\rho$ becomes infinite. It is therefore appropriate to see the full Einstein de Sitter universe as a limiting case of a closed cyclic universe in which the amplitude of the cycle has become infinite as well as the period. Thus, instead of there being infinitely many finite cycles of finite amplitude, there is one infinitely long cycle of infinite amplitude. The infinite cycle is split into two halves of a cycle, the first half being the branch descending from infinity, and the second half, the traditional expanding branch rising to infinity. Since the amplitude of the cycle in the closed universe is proportional to the total mass of the universe as given in (28), the flat Einstein de Sitter universe may be thought of as a limiting case of a closed universe with infinite mass and infinite radius of curvature as discussed in Einstein [23], Silberstein [24], and the author [25]. Also, because of the time-reversibility of the solution, it reads the same, whether one goes from the negative branch to the positive branch or vice-versa. Since, on physical grounds, it is unlikely that before the big bang the universe collapsed in from infinity, but rather collapsed in from a finite size, the cyclic, closed universe seems preferable. The problem of the cusp in $a(t)$ and the singularity in $\rho(t)$ at $t = 0$ can be dealt with classically [26], but a discussion of this lies outside the scope of this work.

Finally, in I it was pointed out that there is a fundamental difficulty with the cosmological term $\Lambda$, based on a generalization of Newton's first law, that led to the conclusion that $\Lambda$ vanishes. It is therefore of interest to briefly recapitulate and extend an argument based on a closed universe that sets an upper bound on $\Lambda$, assumed nonnegative, which would make it too small to lead to the universe accelerating now, as is claimed. The finding arose, nearly a decade before the accelerating universe was proposed, inresponse to work by Weinberg [27] [28], who found, using the weak anthropic principle [29], an upper bound on $\Lambda$ that was based on assuming that the universe is flat, and does not collapse back on itself, that proved to be sufficiently large as to include the current value of $\Lambda$. Following Weinberg's work, the author was able to show that for a closed universe that does collapse back on itself, one obtains a much smaller bound [30]; this was further discussed and extended in [31] which was referenced in I.

To show this here, note that with a cosmological term present, the field Equation (15) for a closed universe becomes

$$\dot{a}^2 - (8\pi G/3)\rho a^2 - \left(\Lambda c^2 a^2/3\right) = -c^2 . \tag{39}$$

Either from the field equation for stress, or by differentiating the above equation, taking into account that $\rho a^3$ is a constant, the equation of motion becomes,

$$\ddot{a} = -(4\pi G/3)\rho a + \left(\Lambda c^2 a/3\right) . \tag{40}$$

Under the assumption that the universe does collapse back on itself after reaching the value of the expansion parameter when it stops increasing and starts to recoil denoted by $a_r$, with $\dot{a}_r = 0$, and $\ddot{a}_r < 0$, it follows from (40) that $\Lambda c^2 < 4\pi G \rho_r$ From the conservation of mass-energy from the contracted Bianchi identities, one has that $\rho_r a_r^3 = \rho_0 a_0^3$, and hence, upon introducing the cosmological term's mass density (rather than energy density) defined as $\rho_\Lambda \equiv \Lambda c^2/8\pi G$, the inequality becomes

$$\rho_\Lambda < (\rho_0/2)(a_0/a_r)^3 . \tag{41}$$

After dividing both sides of the inequality by $\rho_c$, which was not done in [30] [31], one has

$$\Omega_\Lambda < (\Omega_m/2)(a_0/a_r)^3 . \tag{42}$$

Thus, under these circumstances, $\Omega_\Lambda$ would clearly be too small to account for the accelerating universe, for which $\Omega_\Lambda > \Omega_m$. However, because of the above assumptions, for which there is as yet no empirical support, the result is obviously not conclusive, although it is consistent with the rejection of the cosmological term in the proposed alternative to the accelerating universe.

## 5. Concluding Remarks

It is clear from the work in the preceding sections, as well as that in I and II, that the proposed alternative model, based on the slowing down of light by the dark energy in intergalactic space, can explain the diminished brightness of the SNe Ia, and the increased distance to the "standard ruler" of the BAO, as well as can the accelerating universe, that is based on attributing a negative pressure to the dark energy, such as displayed by the cosmological term. However, the crucial test for the alternative model will be for astronomers to determine through suitable observations, such as the one described in Section 4 of II, whether in fact the speed of light in intergalactic space for, say, $z \le 0.7$ is $c/n$, with $n \approx 1.50$. If eventual astronomical observation should show that this is indeed the case, it will then prove theoretically challenging to obtain a general expression for $n$ as a function of redshift, and to show further that $n$ does not exhibit any evidence of dispersion in the optical range, as found by the SNe Ia studies.

Finally, it follows from the discussion in Section 3 that the alternative model is predicting a greater age for the universe than that predicted by the accelerating universe. As was further discussed there, this has bearing on the maximum age of stars, and the value of the Hubble constant, so that their more accurate determination should provide two other areas to test the model astronomically.

## Acknowledgements

I am indebted to Prof. David Weinberg for helpful critical comments concerning the alternative model, and the problems of obtaining astronomical verification. I am also indebted to Prof. Don A. VandenBerg for clarifying the uncertainties in the age of HD 140283, the difficulties in reducing them, and much helpful discussion. I am further indebted to Dr. Gregory B. Malykin for numerous questions, and assistance in obtaining astronomers' comments on the model.

## References

[1] Tangherlini, F.R. (2015) *Journal of Modern Physics*, **6**, 78-87. http://dx.doi.org/10.4236/jmp.2015.61010

[2] Tangherlini, F.R. (2015) *Journal of Modern Physics*, **6**, 1360-1370. http://dx.doi.org/10.4236/jmp.2015.69141

[3] Perlmutter, S., *et al.* (1998) *Nature*, **391**, 51-5. (*Erratum*, **392**, 311.)

[4] Perlmutter, S., *et al.* (1999) *Astrophysical Journal*, **517**, 565-586. http://dx.doi.org/10.1086/307221

[5] Riess, A., *et al.* (1998) *Astronomical Journal*, **116**, 1009-1038. http://dx.doi.org/10.1086/300499

[6] Schmidt, B., *et al.* (1998) *Astrophysical Journal*, **507**, 46-63. http://dx.doi.org/10.1086/306308

[7] Anderson, L., *et al.* (2012) *Monthly Notices of the Royal Astronomical Society*, **427**, 3435-3467. http://dx.doi.org/10.1111/j.1365-2966.2012.22066.x

[8] Anderson, L., *et al.* (2014) *Monthly Notices of the Royal Astronomical Society*, **439**, 83-101. http://dx.doi.org/10.1093/mnras/stt2206

[9] Wang, Y. (2010) Dark Energy. Wiley-VCH, Verlag GmbH & Co, KGaA, Weinheim, 35-49.

[10] Riess, A.G, Nugent, P.E., Gilliland, R.L., Schmidt, B.P., Tonry, J., Aguilar, J.A., *et al.* (2001) *Astrophysical Journal*, **560**, 49-71. http://dx.doi.org/10.1086/322348

[11] Abe, P.A.R., *et al.* (2015) *Astronomy and Astrophysics*, **584**, 1-67. http://dx.doi.org/10.1051/0004-6361/201526586e

[12] Bond, H.E., Nelan, E.P., VandenBerg, D.A., Schaefer, G.H. and Harmer, D. (2013) *Astrophysical Journal Letters*, **765**, L12. http://dx.doi.org/10.1088/2041-8205/765/1/L12

[13] Riess, A. G., *et al.* (2016) Astro-ph, arXiv.1604. 0142442

[14] Riess, A. G., *et al.* (2011) *Astrophysical Journal*, **730**, Article ID: 119. http://dx.doi.org/10.1088/0004-637X/730/2/119

[15] Cheng, C., and Huang, Q. (2015) *Science China Physics, Mechanics and Astronomy*, **58**, 599801. http://dx.doi.org/10.1007/s11433-015-5684-5

[16] Tangherlini, F.R. (1984) *Annals of the New York Academy of Sciences*, **422**, 385-389. http://dx.doi.org/10.1111/j.1749-6632.1984.tb23407.x

[17] Tangherlini, F.R. (1993) *Il Nuovo Cimento B*, **108**, 911-930. http://dx.doi.org/10.1007/BF02828738

[18] Guth, A.H. (1981) *Physical Review D*, **23**, 347-356. http://dx.doi.org/10.1103/PhysRevD.23.347

[19] Guth, A.H. (1997) The Inflationary Universe. Addison-Wesley, Reading, 167-187.

[20] Linde, A.D. (1982) *Physics Letters B*, **108**, 389-392.
http://dx.doi.org/10.1016/0370-2693(82)91219-9

[21] Albrecht, A. and Steinhardt, P.J. (1982) *Physical Review Letters*, **48**, 1220-1223.
http://dx.doi.org/10.1103/PhysRevLett.48.1220

[22] Tangherlini, F.R. (1993) *Il Nuovo Cimento B*, **108**, 1253-1273.
http://dx.doi.org/10.1007/BF02741278

[23] Einstein, A. (1923) The Meaning of Relativity. Translation by Adams, E.P., 3rd Edition, Princeton University Press, Princeton, 107.

[24] Silberstein, L. (1924) The Theory of Relativity. 2nd Edition, Macmillan, London, 479-480.

[25] Tangherlini, F.R. (1997) *General Relativity and Gravitation*, **29**, 869-880.
http://dx.doi.org/10.1023/A:1018835300818

[26] Tangherlini, F.R. (1998) *Il Nuovo Cimento B*, **113**, 785-804.

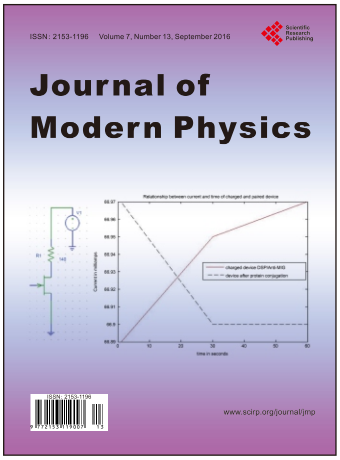
[27] Weinberg, S. (1987) *Physical Review Letters*, **59**, 2607-2610.
http://dx.doi.org/10.1103/PhysRevLett.59.2607

[28] Weinberg, S. (1989) *Reviews of Modern Physics*, **61**, 1-23.
http://dx.doi.org/10.1103/RevModPhys.61.1

[29] Barrow, J.D. and Tipler, F. (1986) The Anthropic Cosmological Principle. Oxford University Press, New York, 15-26.

[30] Tangherlini, F.R. (1989) *Il Nuovo Cimento B*, **103**, 311-317.
http://dx.doi.org/10.1007/BF02737134

[31] Tangherlini, F.R. (1991) *Il Nuovo Cimento B*, **106**, 123-145.




# Call for Papers

# Journal of Modern Physics

**ISSN: 2153-1196 (Print)   ISSN: 2153-120X (Online)**
**http://www.scirp.org/journal/jmp**

***Journal of Modern Physics*** **(JMP)** is an international journal dedicated to the latest advancement of modern physics. The goal of this journal is to provide a platform for scientists and academicians all over the world to promote, share, and discuss various new issues and developments in different areas of modern physics.

## Editor-in-Chief

**Prof. Yang-Hui He**   City University, UK

## Executive Editor-in-Chief

**Prof. Marko Markov**   Research International, Buffalo Office, USA

## Subject Coverage

Journal of Modern Physics publishes original papers including but not limited to the following fields:

- Biophysics and Medical Physics
- Complex Systems Physics
- Computational Physics
- Condensed Matter Physics
- Cosmology and Early Universe
- Earth and Planetary Sciences
- General Relativity
- High Energy Astrophysics
- High Energy/Accelerator Physics
- Instrumentation and Measurement
- Interdisciplinary Physics
- Materials Sciences and Technology
- Mathematical Physics
- Mechanical Response of Solids and Structures
- New Materials: Micro and Nano-Mechanics and Homogeneization
- Non-Equilibrium Thermodynamics and Statistical Mechanics
- Nuclear Science and Engineering
- Optics
- Physics of Nanostructures
- Plasma Physics
- Quantum Mechanical Developments
- Quantum Theory
- Relativistic Astrophysics
- String Theory
- Superconducting Physics
- Theoretical High Energy Physics
- Thermology

We are also interested in: 1) Short Reports—2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data; 2) Book Reviews—Comments and critiques.

## Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

## Website and E-Mail

**http://www.scirp.org/journal/jmp**   **E-mail: jmp@scirp.org**