The Journal on Advanced Studies in Theoretical and Experimental Physics,
including Related Themes from Mathematics

PROGRESS IN PHYSICS

"All scientists shall have the right to present their scientific research results, in whole or in part, at relevant scientific conferences, and to publish the same in printed scientific journals, electronic archives, and any other media." — Declaration of Academic Freedom, Article 8

This journal is powered by LaTeX

A variety of books can be downloaded free from the Digital Library of Science:
http://fs.gallup.unm.edu/ScienceLibrary.htm

April 2020   Vol. 16, Issue 1

CONTENTS

Information for Authors

Progress in Physics has been created for rapid publications on advanced studies in theoretical and experimental physics, including related themes from mathematics and astronomy. All submitted papers should be professional, in good English, containing a brief review of a problem and obtained results.

All submissions should be designed in LaTeX format using *Progress in Physics* template. This template can be downloaded from *Progress in Physics* home page http://www.ptep-online.com

Preliminary, authors may submit papers in PDF format. If the paper is accepted, authors can manage LaTeXtyping. Do not send MS Word documents, please: we do not use this software, so unable to read this file format. Incorrectly formatted papers (i.e. not LaTeXwith the template) will not be accepted for publication. Those authors who are unable to prepare their submissions in LaTeXformat can apply to a third-party payable service for LaTeX typing. Our personnel work voluntarily. Authors must assist by conforming to this policy, to make the publication process as easy and fast as possible.

Abstract and the necessary information about author(s) should be included into the papers. To submit a paper, mail the file(s) to the Editor-in-Chief.

All submitted papers should be as brief as possible. Short articles are preferable. Large papers can also be considered. Letters related to the publications in the journal or to the events among the science community can be applied to the section *Letters to Progress in Physics*.

All that has been accepted for the online issue of *Progress in Physics* is printed in the paper version of the journal. To order printed issues, contact the Editors.

Authors retain their rights to use their papers published in *Progress in Physics* as a whole or any part of it in any other publications and in any way they see fit. This copyright agreement shall remain valid even if the authors transfer copyright of their published papers to another party.

Electronic copies of all papers published in *Progress in Physics* are available for free download, copying, and re-distribution, according to the copyright agreement printed on the titlepage of each issue of the journal. This copyright agreement follows the *Budapest Open Initiative* and the *Creative Commons Attribution-Noncommercial-No Derivative Works 2.5 License* declaring that electronic copies of such books and journals should always be accessed for reading, download, and copying for any person, and free of charge.

Consideration and review process does not require any payment from the side of the submitters. Nevertheless the authors of accepted papers are requested to pay the page charges. *Progress in Physics* is a non-profit/academic journal: money collected from the authors cover the cost of printing and distribution of the annual volumes of the journal along the major academic/university libraries of the world. (Look for the current author fee in the online version of *Progress in Physics*.)

————————

The Interpretation of the Hubble-Effect and of Human Vision Based on the Differentiated Structure of Space

Gerhard Dorda

Institute of Physics, University of Armed Forces Muenchen, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany.
E-mail: physik@unibw.de

Based on the differentiated structure of space, observed by the Quantum-Hall-Effect, a comprehensive equation is presented for the description of the Hubble-Effect. This Hubble-Effect equation reflects the experimental observation showing a casual connection to the Hubble time T_U and thus to the cosmic length L_U and the cosmic mass M_U. The obtained results are substantiated by the cosmic background radiation and by the agreement of the derived data with the experimental data of the Milky Way. It is shown that the differentiated structure of space, used for the description of the Hubble-Effect, also refers to the process of human vision, dominating the observation.

1 Introduction

After the discovery of the Quantum-Hall-Effect (QHE) and the associated exceptional side effects [1], it proved to be necessary to re-evaluate many physical and biological phenomena, e.g. the interpretation of the Hubble-Effect (HE) and, of the basis of it, even the process of human vision, referring to the *differentiated structure of space*. The differentiated structure of the three-dimensional space was first observed at the analysis of the experimental data of the QHE, which was discovered in 1980 by K. von Klitzing based on MOS-field-effect transistors.

The QHE is the first experimental observation of *quantization in the macroscopic scale* in solid-state physics. Only gradually, the fundamental importance of this discovery and of all with this discovery connected spectacular experimental observations became apparent for the entire range of physics. In the first instance, it was the observation of the QHE on GaAs–Al$_x$Ga$_{(1−x)}$As heterostructures [2], presented by D. C. Tsui et al, which showed that this effect is generally valid for the whole solid-state physics. More detailed investigations of the experimental data revealed that the QHE is not only independent from atomic mass, but also from the strength of the electric current used, i.e. from frequency, i.e. from time, and also from the form of the sample with the considered QHE structure, i.e. from space [3].

Really, the state of QHE shows a spectacular simultaneity of $R_{xx} = h/ie^2 = 2.58128 \times 10^4/i\,\Omega$ *and* $R_{xx} = 0\,\Omega$ (i is the quantization number), measured between different contacts *at any place* of the QHE structure. This effect of the spatial independence of the observed simultaneity in resistivity is the background of the disclosed two-dimensionality of electromagnetism at the causal situation. Besides that, it should be emphasized that the simultaneity of the quantized resistivity shows that the three-dimensional state of electromagnetism can be clearly separated spatially in two independent conditions: On one side in a 2-D state, given by the simultaneity, and on the other side in a 1-D state, realized capacitively by

the interaction of the electron charges. The experimental observation of the possibility to split up electromagnetism in a 2-D and a 1-D state will be described by the "differentiated structure of the space" [3]. Analyzing all these novel experimental insights allowed to deliver convincing physical answers, for example Lee Smolin's book *The Trouble with Physics* posed fundamental and unsolved questions [4], in particular also about the category of time [3].

The description of space and time, i.e. frequency, based on the QHE, leads to the notion that also open questions in astronomy and cosmology could be answered with the help of the observations of the QHE. This, for example, includes the question about cosmic expansion, which, on the basis of the interpretation of the Hubble-Effect (HE), generated a vivid discussion, leading to the unfolding of several cosmic models, but without final solutions [5, 6]. Therefore, in this work, it is attempted to explore the experimental data of the HE on the basis of the so-called differentiated structure of space [3].

2 The analysis of the Hubble-Effect (HE) with respect to the differentiated structure of space

The cosmic expansion model is based on the experimentally observed Hubble-law, given by [5]

$$v_{HE,y} = \frac{R_{HE,y}}{T_U} \,. \tag{1}$$

Here in (1), $v_{HE,y}$ is the velocity of a given galaxy, $R_{HE,y}$ has the significance of a distance referred to a given galaxy and T_U is interpreted as the Hubble time, defining the so-called age of the cosmos (an assumption which requires the expansion of the cosmos). The index $_{HE}$ signifies the relation of the Hubble-Effect (HE) to the associated redshift of the observed radiation and the index $_y$ refers this redshift to the observed galaxy [5, 6].

The figures of the experimental HE in [5] and [6] show the so-called escape velocity $v_{HE,y}$ in relation to the velocity of light c, meaning that (1) can be rewritten by use of c. As a result, we receive a form which defines the HE in relation

to the so-called length of the cosmos, obtained by $L_U = T_U c$, and we may write

$$\frac{R_{HE,y}}{L_U} = \frac{v_{HE,y}}{c} \, . \tag{2}$$

The value of the redshift is usually specified by the number z_y, which means

$$z_y = \frac{v_{HE,y}}{c} \, . \tag{3}$$

Since the HE merely reflects the observation of light, i.e. photon energies, the number z_y may, in accordance with (1) and (2) and due to the c-standardization, be considered to be related to the limit of the light frequency f_C or to the limit of the light wavelength λ_C. As shown in Section 4, this is of fundamental importance for the interpretation of the HE.

The concept of an escape velocity $v_{HE,y}$, as stated in (1), must originate from the existence of a given position, e.g. from the place of observation, or in a general sense from any localized place in the cosmos, in order to have the possibility to speak of place in sense of the classic conception of velocity, a model, which so far has been crucial for the interpretation of the HE. The concept of a place requires the existence of localization related to atomic mass, i.e. to protons and neutrons, constituting a gravitationally induced localization which only can become real through an atomic solid-state structure.

Starting from these findings it can be shown that based on the experimental data of the QHE, which is independent of atomic mass, a novel form of velocity can be defined. This velocity is also given by the relation of length and frequency, but this specific form of velocity is merely deduced from the dualistic character of the electron, i.e. without any contribution of proton-neutron-mass related gravity. This specific i.e. structural space-time condition, which is identifiable in the QHE, reveals that the electron-related velocity is given by the relation of the category of length, reflected by the electron mass m_e, and the category of frequency, realized by two-dimensional electromagnetism, i.e. by the electron charge e. This length-frequency, i.e. length-time relation is, in spatial terms, always mutually perpendicular to each other, which is the background for the notion of three-dimensionality of space and also the background for the *freedom* of choice concerning the value of light velocity. As shown in [3, pp. 33–34,45,49–50], it therefore follows the possibility of differentiation between the one-dimensionality, i.e. 1-D, and the two-dimensionality, i.e. 2-D. These fundamental circumstances were characterized in summary as a differentiated three-dimensional spatial structure.

It is evident that this electron related form of velocity is given at light effects, i.e. given by λf (λ = wavelength, f = frequency). Thus, it can be assumed that this form of velocity is also displayed in the observation of the HE-galaxies, playing an essential key role in the here presented reinterpre-

tation of the HE. Hence, unexpected statements about the HE-galaxies may be obtained when the Hubble-law, i.e. (1), and the model of the differentiated structure of space are applied to Kepler's third law.

3 The application of the Hubble-Effect to Kepler's third law

To begin with, it seems necessary to appropriately transform Kepler's third law. In doing so, we assume that due to the cosmological principle [6], Kepler's third law has general validity in the entire universe.

Kepler's third law is given by [3],

$$\left(\frac{T_{G,y}}{2\pi}\right)^2 = t_{G,y}^2 = \frac{R_{G,y}^3}{G \, M_{G,y}} \, , \tag{4}$$

whereby G in (4) is the gravitational constant, given by

$$G = c^2 \frac{L}{M} \, . \tag{5}$$

Eq. (4), in conformity with the MKSA- or MKS-system of units, represents a universal linkage of the category of length with the category of time, modified by the category of mass. $T_{G,y}$ in (4) is the so-called orbital period of the given solid-state celestial body (SSCB), which planets, suns and stars are to be counted as part of. $t_{G,y}$ in (4) is the so-called effective time, referred to the surface of the SSCB, $R_{G,y}$ is the distance to the center of the SSCB and $M_{G,y}$ its mass. The index G signifies the connection to the SSCBs. In (5), L bears the meaning of the Planck length, $L = 4.051 \times 10^{-35}$ m, and M represents the Planck mass, $M = 5.456 \times 10^{-8}$ kg [7]. By transforming (4), we receive the following form, being valid for all SSCBs

$$\frac{v_{G,y}^2}{c^2} = \frac{L}{R_{G,y,1-D}} \frac{M_{G,y}}{M} \, , \tag{6}$$

whereby

$$v_{G,y} = \frac{R_{G,y,2-D}}{t_{G,y}} \, . \tag{7}$$

In (6), the left-hand side represents the electromagnetic effect, i.e. an effect reflecting spatial two-dimensionality, and the right-hand side reflects a distance related, i.e. a one-dimensionality related gravitational effect.

Here, in (6) and (7), the findings from the Quantum-Hall-Effect (QHE) about the possibility of the differentiated space is used, according to which the three-dimensional space, in case of it being structured, can be considered partitioned, and that [3]:

1. in a one-dimensional space, described by the 1-D state, covered by $R_{G,y,1-D}$, and

2. in a two-dimensional space, described by the 2-D state, ascertainable by $R_{G,y,2-D}^2$.

Attention should be paid to the fact that the one-dimensional gravitational distance $R_{G,y,1-D}$ of the SSCBs, as given in (6), could be described by the number $a_{G,y}$, which due to the reference to one-dimensionality was termed gravitational number. In [3, see p. 14], it is given by

$$R_{G,y,1-D} = a_{G,y}\,\lambda_{G,y}\,. \tag{8}$$

Here, $\lambda_{G,y}$ is a one-dimensional reference length, defined by

$$\lambda_{G,y} = M_{G,y}\frac{L}{M}\,. \tag{9}$$

It is easily recognizable that in accordance with (4) – (8), this reference length $\lambda_{G,y}$ signifies the connection between the category length and the atomic mass related gravitation.

When discussing (6), it is of importance to consider that the Planck relation L/M in (5) and (9) possess, due to the cosmological principle, validity for the entire being in the cosmos. Therefore, as an extension of L/M, we may write

$$\frac{L}{M} = \frac{\lambda_{G,y}}{M_{G,y}} = \frac{L_U}{M_U}\,, \tag{10}$$

which is a consequence of the general validity of Kepler's third law. Here in (10), $\lambda_{G,y}$ stands for the reference length of the SSCB and $M_{G,y}$ for its related mass. Furthermore, L_U and M_U are the limit length L_U and the limit mass M_U of the cosmos, introduced by means of (1) and (2), i.e. by means of the HE.

The masses $M_{G,y}$ in (4), (6), and (9) are effective as homogeneity parameters. As will be shown, the state of homogeneity can be related to two different structures in the cosmos, which, in the three-dimensional cosmic space, are identifiable by their dot-like centered unity. These two forms are:

1. Celestial bodies which consist of solid state, i.e. SSCBs, and which can, by means of Kepler's third law, be very well described as spherical structures, given by interwoven gravitational-electromagnetic structures ([3], page 44). All planets, suns and stars are to be counted as part of this. With regard to (4), the boundary condition for the homogeneity of the SSCB is the equality $R_{G,y,1-D} = \sqrt{R_{G,y,2-D}^2}$, which enables dynamics, i.e. the category of time, to be revealed in Section 4.

2. Celestial bodies whose existence only is observable with the aid of optical methods, i.e. with the aid of eyesight and technically with the aid of optical absorption methods. This includes galaxies, theoretically ascertained by (1) and (2) of the HE. These cosmic structures are not given by a coherent, gravitational-electromagnetic interwoven state, but they are to be considered a free, i.e. dynamic cluster of different SSCBs, which, as part of above all electromagnetic interactions, form

by the so-called "black hole" a homogeneous, i.e. dot-like centered unity. Due to the free cluster of SSCBs, which show only insignificant gravitational interaction, the possibility of creating the category of time by means of galaxies does not exist. Hence, we are able to clarify the boundary condition for the homogeneity of the HE-galaxies only in Section 5.

To clearly show the difference between the SSCBs and the galaxies, (6) must be adapted to (1) and (2). Based on (6) and (10), we may write

$$\left(\frac{v_{HE,y}}{c}\right)^2 = \frac{L_U}{R_{HE,y,1-D}}\frac{M_{HE,y}}{M_U}\,, \tag{11}$$

whereby $v_{HE,y}$ is given by

$$v_{HE,y} = \frac{R_{HE,y,2-D}}{T_U}\,. \tag{12}$$

$M_{HE,y}$ signifies the mass related to the given galaxy. The distances $R_{HE,y,1-D}$ and $R_{HE,y,2-D}$ in (11) and (12) are, according to the cosmological principle, to be interpreted as characteristic distances, i.e. lengths, of the given galaxy.

In conformity with (6) and (11), the fundamental difference between the SSCBs and the HE-galaxies should become above all apparent by means of the different definitions of $v_{G,y}$, (7), and of $v_{HE,y}$, (12). Thus, this difference is discussed in the following sections.

4 The difference between solid-state celestial bodies (SSCBs) and HE-galaxies

When comparing the velocities $v_{G,y}$ and $v_{HE,y}$, we proceed that both $R_{G,y,2-D}$, the distance of the given solid-state celestial body (SSCB), and $R_{HE,y,2-D}$, the distance of the given galaxy, are to be considered their distinctive characteristic. In doing so, the cosmological principle is to be heeded, stating that in the cosmos there is no center and consequently no defined position [6]. Moreover, the fundamental difference between the time statements $t_{G,y}$ and T_U, given in (4) and (1), has to be taken into account since it points out that, as (4) and (6) show, the time $t_{G,y}$ is one of the characteristic parameters of any given SSCB, whereas the time T_U, being valid for all HE-galaxies, is solely a cosmic constant. From Kepler's third law, (4) and (6), it results that the time $t_{G,y}$ is given by

$$t_{G,y} = \sqrt{a_{G,y}}\,\frac{R_{G,y,2-D}}{c}\,, \tag{13}$$

whereby $a_{G,y}$ is the SSCB related gravitational number, defined in (8). Thus, considering (6), (7), and (13), the solid-state celestial body is characterized not only by the mass $M_{G,y}$ and the radius $R_{G,y}$, but also by the SSCB related category of time $t_{G,y}$.

In contrast to $v_{G,y}$, the velocity $v_{HE,y}$ can experimentally only be experienced by optical means, in fact with aid of the

light i.e. photon energies, emitted by the given galaxy. This energy spreads from the galaxy with the velocity of light and is registered by the eye or by appropriate appliances (telescopes) via absorption. Since the respective galaxies distinguish from each other by the emitted light i.e. photon energy, it is physically permitted, in compliance with the observed value of the so-called redshift z_y, to ascribe an appropriate frequency f_y to the observed galaxy, which reflects the energy $h f_y$. That means, the in (2) presented relation $v_{HE,y}/c$ can be replaced by an appropriate frequency or wavelength relation, and we may write

$$z_y = \frac{h f_y}{h f_C} = \frac{hc/\lambda_y}{hc/\lambda_C} = \frac{\lambda_C}{\lambda_y} \, . \tag{14}$$

It then again follows that the HE can be described by means of an equation of light

$$\lambda_y f_y = \lambda_C f_C = c \, , \tag{15}$$

which inter alia reflects the fact that the frequency, in localized form known as the category of time, is an expression of pure electromagnetism [3].

Here in (14) and (15), f_y is the given galaxy related frequency or λ_y wavelength, whereas f_C is the Compton frequency and λ_C the Compton wavelength. In (14), z_y is, unlike in the classic Doppler-effect model, not valued as a difference from wavelengths, but as a direct information about the observed galaxy state, given by f_y or λ_y, respectively. Thus, (14) and (15) determine the state of the HE-galaxies. Hence, instead of interpreting $v_{HE,y}$ mechanically as an escape velocity of the galaxies, it proves to be physically acceptable, with regard to (14) and (15), to replace the concept of the classical velocity with the frequency or wavelength relation given by (14) and to describe the redshift as a light wave radiation, which reflects the heat radiation laws, i.e. Wien's displacement law. That means, it is postulated that any HE-galaxy emits radiation in the form of photon energy as a result of its homogeneity.

5 The equation of the Hubble-Effect

Starting the analysis of this novel description of the HE, above all it must be emphasized that the existence of the parameter of the HE galaxies, given by T_U, attests the validity of Kepler's third law for the whole cosmos, i.e. the form of the gravitational constant (5), and also the extension of L/M, presented in (10). Thus it is – from a physical point of view – legitimate to use the (4), (5), and (10) as basic equations for the further analysis of (11), at which we take the form $R_{HE,y,1-D}$ in place of $\lambda_{G,y}$ of (9). Furthermore, it appears absolute necessary for the description of the HE to apply the model of the differentiated structure of the space to (11). This requirement indicates to formulate (11) in a particular form, reflecting this spatial differentiation. It can be achieved by a completion of

(11) by the factor z_y^2, hence formulating

$$\left(\frac{v_{HE,y}}{c} \right)^2 = \frac{L_U \, z_y^2}{R_{HE,y,1-D}} \frac{M_{HE,y}}{M_U} \, . \tag{16}$$

Really, it should be considered, the experimental HE data shows that the factor z_y is causally related to the distance $R_{HE,y,2-D}$, as it was on the basis of (2) and (3) expressed by (12). Thus, to be in accordance with the required differentiation of the HE-state from the usual three-dimensionality into the one-dimensionality and the two-dimensionality, we have to conclude that the factor z_y^2 must be related to the 1-D related distance $R_{HE,y,1-D}$, to ensure the causality at the whole HE-state. Evidently, these requirements are realized by means of (16).

Taking into consideration all the presented experimental data and the related conclusions, given in Sections 2 – 4, we are able to present the solution of the whole HE state, and that in form of a comprehensive, generally valid equation, given by

$$\begin{aligned} R_{HE,y,1-D} &= z_y \, R_{HE,y,2-D} = z_y^2 \, L_U \\ &= z_y^x \, M_{HE,y} = z_y^{x+2} \, M_U \, , \end{aligned} \tag{17}$$

at which z_y^x is given by

$$z_y^x = \frac{L}{M} = \frac{R_{HE,y,1-D}}{M_{HE,y}} = \frac{L_U}{M_U} = 7.426 \times 10^{-28} \, \text{m kg}^{-1} \, . \tag{18}$$

Equation (17) shows that with respect to (2) it is possible to formulate the relations

$$R_{HE,y,2-D} = z_y \, L_U \, , \tag{19}$$

as well as

$$R_{HE,y,1-D} = z_y \, R_{HE,y,2-D} \tag{20}$$

and

$$M_{HE,y} = z_y^2 \, M_U \, . \tag{21}$$

Furthermore, on the basis of (16), it becomes evident that the difference between the SSCBs and the state of HE-galaxies is simply describable by the factor z_y, which is, according to (6), for the SSCBs without exception given by $z_y = 1$.

The numerical value of (18) results from the experimentally explored gravitational constant [4] $G = 6.6738 \times 10^{-11}$ m^{3}kg^{-1}s^{-2}, using (5). It demonstrates the value z_y^x to be a natural constant. Besides, it should also be emphasized that (18) is therefore significant for our model, as it discloses the functional background of the homogeneity of the HE galaxies.

The validity of (17) and (18), and thus of (19) − (21), can be verified using both the knowledge of the cosmic background radiation and the known experimental data of the Milky Way, since according to our model the Milky Way galaxy is assessed to be a homogeneous galaxy.

6 The analysis of the cosmic background radiation with respect to the Milky Way galaxy radiation

At first, when analyzing (17) and (18), which describe the state of all HE galaxies, it must be pointed out that in the spatially differentiated state, as it is the case for the optical observation of the HE galaxies, only the electron related electromagnetic variability is ascertained, so that a specific proton-neutron one-dimensional mass-effect cannot be observed at this effect by experiment. Thus to solve this problem, the observation of the cosmic background radiation is considered. It shows that this radiation, represented by the temperature $T_{\text{cosm}} = 2.73\,°K$, is the result of the interaction of hydrogen atoms, extended over the whole cosmos, see [5] and [12].

Thus when we value the cosmic background radiation as a heat radiation effect, given by the displacement law of Wien, obtaining $z_T = \lambda_C/\lambda_{\text{cosm}} = T_{\text{cosm}}\,\lambda_C/(3.40 \times 10^{-3})$ [3, part III] and assess this value with respect to the heat radiation factor of the Milky Way, given by $z_{\text{MW}} = \lambda_C/\lambda_{\text{MW}}$, evidently this z_T value has to be modified by the relation m_p/m_e, corresponding to the temperature relation $T_{\text{MW}}/T_{\text{cosm}}$. Here m_p is the mass of the proton, m_e the mass of the electron and T_{MW} has reference to z_{MW}. In other words, the factor of modification m_p/m_e represents the energetic difference between the cosmic background radiation, being a result of the interaction of hydrogen atoms, and the radiation of the localized, i.e. spatially differentiated electromagnetism of the HE galaxies.

Using the HE-related (9) and (10), as well as the (17) and (20), and assuming that the heat radiation factor of the Milky Way is identical with the cosmic background radiation factor z_T, then we obtain the following relationship

$$R_{\text{HE,MW,2–D}} = \frac{3.4 \times 10^{-3}}{T_{\text{cosm}}\,\lambda_C}\frac{m_e}{m_p}\frac{L_U}{M_U}M_{\text{HE,MW}} \tag{22}$$
$$= 2.08 \times 10^{-22}\,M_{\text{HE,MW}}\,.$$

Here, in place of the Milky Way radiation factor z_{MW}, the assumed identity of z_{MW} to z_T was used, resulting in

$$z_{\text{MW}} = z_T = \frac{T_{\text{cosm}}\,\lambda_C}{3.4 \times 10^{-3}}\frac{m_p}{m_e} = 3.58 \times 10^{-6}\,. \tag{23}$$

At (23), T_{cosm} was replaced with the background radiation value $T_{\text{cosm}} = 2.73\,°K$, and the causal relation $\lambda_{\text{max}} T = 3.40 \times 10^{-3}$, i.e. Wien's displacement law, was used for λT. When we use the Hubble time $T_U = 4.32 \times 10^{17}$ s, reflecting a Hubble constant of $H_o = 71.4\,\text{km s}^{-1}\text{Mpc}^{-1}$, lastly obtained from the Hubble telescope, we obtain a cosmic length $L_U = c\,T_U = 1.30 \times 10^{26}$ m, and by means of (10) a cosmic mass $M_U = 1.74 \times 10^{53}$ kg.

Finally, by means of $(19)–(23)$, for the Milky Way we obtain the values

$$R_{\text{HE,MW,2–D}} = z_{\text{MW}}\,L_U = 15.05\,\text{kpc} = 4.63 \times 10^{20}\,\text{m}\,,$$

$$M_{\text{HE,MW}} = z_{\text{MW}}^2\,M_U = 1.12 \times 10^{12}\,\text{solar masses} \tag{24}$$
$$= 2.23 \times 10^{42}\,\text{kg}\,.$$

Considering these results with respect to the experimentally observed data of the Milky Way, given in [6] by the approximate values of the radius $R_{\text{HE,MW,2–D}} = 15\,\text{kpc} = 4.6 \times 10^{20}$ m and of the mass $M_{\text{HE,MW}} = 10^{12}$ solar masses $= 2 \times 10^{42}$ kg, we assess for the radius a factor of inaccuracy of only 3 %, and for the mass $M_{\text{HE,MW}}$ of only 12%. This finding, especially the agreement of the order of magnitude of both $R_{\text{HE,MW,2–D}}$ and $M_{\text{HE,MW}}$, is very important, as it convincingly demonstrates that (17) and (18) can be assessed as a novel, physically justified description of the Hubble-effect.

The Sections 2–6 have shown that the novel HE model is based on the QHE-observation about the differentiated structure of the 3-dimensional space. The application of the differentiated space structure on the gravitational constant G , (5), shows that c^2 is related to electromagnetism, in the case of the HE-galaxy to the 2-D state, represented by the $R_{\text{HE,y,2–D}}$ distance, whereas L/M refers to the gravity of the HE-galaxies, i.e. to the 1-D state, represented by the $R_{\text{HE,y,1–D}}$ distance, which is in this situation in a causal connection to the mass $M_{\text{HE,y}}$. The HE-circumstance, described by $(16)–(18)$ and thus by (20), shows that the connection between the distances $R_{\text{HE,y,2–D}}$ and $R_{\text{HE,y,1–D}}$ is given by the factor z_y.

These results are confirmed by the agreement of the calculated data with the experimental data of the Milky Way and support also the conception, formulated by (10), that the relation $R_{\text{HE,y,(1–D)}}/M_{\text{HE,y}}$ of *any* HE galaxy is always identical with the L/M-relation.

7 The description of human vision on the basis of the differentiated structure of space

A particular confirmation of the value z_{MW} is obtained by considering the general limitation of vision. Seen in this connection, it should be pointed out that not only that of human eyes, but also the vision of all animals breaks off at the wavelength $\lambda_y = 6.8 \times 10^2$ nm [13]. This particular observation manifests the rightness of the identity between the limiting value of the wavelength of visible light and the specific wavelength of the radiation of the Milky Way $\lambda_{\text{MW}} = 6.79 \times 10^2$ nm.

A further very interesting observation about the process of seeing is obtainable, when we become aware of the connection between human vision and the differentiated structure of space. As disclosed extensively in *The Feynman Lectures on Physics* [13], human vision is the result of processing of two signals, independently given on the one side by the rod cells, and on the other side by the uvula cells. In this textbook, it is shown that the rod cells yield signals at the twilight, i.e. signals without any colored light absorption, whereas the uvulas show signals solely by means of colorful light.

This biological differentiation reflects in an absolute manner the physical model of spatial differentiation between gravitation and electromagnetism, suggesting that the rod-signals represent the 1-D related gravitational interaction, whereas the uvula-signals the 2-D related electromagnetic interaction.

Thus it is physically acceptable to suggest the biological structure of human eyes to be the consequence of the effect of the discussed existence of the differentiated structure of space, given outside of masses. Consequently we can state that this interesting biological differentiation between rods and uvulas reflects the spatial differentiation between the 1-D state and the 2-D state, showing that the differentiated structure of space is the particular mediator of these effects.

Considering these circumstances, it becomes evident that due to the existence of the differentiated structure of space, the human eyes become the main processing not only of the perceptibility of solids and thus of the observation in general, but also, simply by the absorption of particular quanta of light, of the perceptibility of stars and galaxies, and attain therefore, together with the help of telescopes, the possibility to discover the HE and the related equations (16), (17) and (18).

8 Concluding findings

In the cosmos, there are two forms of homogeneous structures: Solid-state celestial bodies (SSCBs) and HE-galaxies (HE). Homogeneous solid-state celestial bodies consist of electromagnetic-gravitational interwoven structures, which can be described by Kepler's third law. This law shows that the SSCBs can not only be characterized by mass and radius, but also by the category time, whose lapse is dependent on the strength of gravity of the given SSCB [3].

In contrast, the existence of HE-galaxies is solely observable by means of optical signals, i.e. by eyes and/or by technical methods, using telescopes. Here, signals undisturbed by atmospheric absorption are required, which correspond to the state of a differentiated three-dimensional space. Incidentally, in this connection it should be emphasized that the Pythagorean theorem, considered in conjunction with the three-body problem, entirely corresponds to this differentiated three-dimensional space model. Therefore, it should be pointed out that the application of the differentiated structure of space to the optical signals of galaxies leads, with analyzing the HE, to (16) − (18). In addition, it was demonstrated that the validity of (17) and (18) can be established by the cosmic background radiation, and what is more, by the excellent agreement of the deduced data of mass and radius of the Milky Way with the corresponding values.

The presented new model of the Hubble-effect, which is based on the black-body radiation, shows – according to the *experimental*, generally valid disclosures of the Quantum-Hall-Effect (QHE) – that the so-called HE-velocity $v_{HE,y}$ is a pure electron effect. Therefore it has been stated that, according to the differentiated structure of space, the frequency, i.e. the category time, should not be considered an absolute basic magnitude, but an electromagnetic 2-D state, which becomes localized, i.e. observable only in connection with the existence of masses. Therefore, as generally known – and

also being in agreement with the differentiated space model – time can be observed only in a causal relation to the 1-D length state [4]. This conclusion follows from (4) and (8) and has been manifested by experimental data of the lapse of time, in particular described in [3] by (26).

Finally, the importance of the differentiated structure of space in nature has been further made evident by the analysis of human vision, showing that the difference of the function of the uvula cells and the rod cells reflects the separateness of the 2-D and 1-D spatial state of seeing, an effect, being in accordance with the description of the Hubble-Effect. This observation is of extraordinary importance, as the process of seeing is the main background of the human observation of all being. As will be shown in a next paper, this important conclusion can be additionally substantiated by the physical description of the process of hearing [14, 15].

References

1. von Klitzing K., Dorda G. and Pepper M. New method for high-accuracy determination of fine-structure constant based on quantized Hall resistance. *Phys. Rev. Lett.*, 1980, v. 45, 494–497.

2. Tsui D. C., Störmer H. L. and Gossard A. C. *Phys. Rev. Lett.*, 1982, v. 48, 1559–1561.

3. Dorda G. Die Struktur von Raum und Zeit und die Interpretation der Wärme. Cuvillier Verlag, Göttingen, ISBN 978-3-7369-9388-4, eISBN 978-3-7369-8388-5, 2016.

4. Smolin L. The Trouble with Physics. Houghton Mifflin Company, New York, 2006. In German: Die Zukunft der Physik. Deutsche Verlags-Anstalt, München, 2009, pp. 344–346.

5. Silk J. A short history of the universe. *Scientific American Library*, W. H. Freeman and Company, New York, 1994. In German: Die Geschichte des Kosmos. Spektrum, Berlin, 1996, pp. 34–42.

6. Bernhard H., Lindner K. and Schukowski M. Wissensspeicher Astronomie. Volk und Wissen Verlag, Berlin, ISBN 3-06-081705-7, 1995.

7. Kostro L. De Broglie waves and natural units. In: van der Merwe A. and Garuccio A., eds. Waves and particles in light and matter. Plenum Press, New York, 1994, 345.

8. Gerthsen C. and Vogel H. Physik. Springer-Verlag, Berlin, 1993, pp. 570–571.

9. Hausamann D. and Further J. Einstein und die Satellitennavigation, In: Einsteins Relativitätstheorien, Deutsches Museum. Deutsches Museum-Verlag, München, 2005, pp. 24–34.

10. Pound R. V. and Rebka G. A. *Phys. Rev. Lett.*, 1960, v. 4, 337–341.

11. Bublath J. Geheimnisse unseres Universums – Zeitreisen, Quantenwelten, Weltformel. Droemer Verlag, München, 1999, ch. 10, pp. 114–115.

12. Fritzsch H. Vom Urknall zum Zerfall. In: Serie Piper SP518, R.Piper & Co. Verlag, München, 1987, pp. 255–256, 283.

13. Feynman R. P., Leighton R. B. and Sands M. The Feynman Lectures on Physics, Vol. 1. California Institute of Technology, 1963, ch. 35–36. In German: Vorlesungen über Physik, Band 1. Oldenbourg Verlag, München, 1987, pp. 477–505.

14. Fletcher H. and Munson W. A. Loudness, its definition, measurement and calculation. *J. Acoust. Soc. Am.*, 1933, v. 5, 82.

15. Dorda G. Die Interpretation des Schalls auf der Basis der differenzierten Struktur des drei-dimensionalen Raumes. University of the Armed Forces, Muenchen, Germany, to be published.

—————————

Predictability Is Fundamental

Felix Tselnik

E-mail: tselnik@bgu.ac.il

1 General concept of trajectory

In his relationship with Nature, the person might be active if he wants to get to some state of the world, and then he is looking for a means to reach this state. Although the content of the state is completely in his mind, he needs the prediction for his action to reach the desired. Typically, this is difficult in real life, and people act according to more or less uncertain hopes, past experience, beliefs etc. However, sometimes predictions might exist to recommend actions with the universally guarantied results – always and everywhere. Though infrequent, such predictions are therefore recommended to be looked for first of all, and our so valued technologies are based solely on these.

The related scheme of the world states must be able to formulate predictions in its own internal terms. If some state in the scheme is associated with the desired, so being the final for the person's purpose, the initial state, from which the action should start, must be defined in the internal terms of the scheme as well. Since the final state is not reached as yet, it should be set in the future with respect to the initial. If being in the initial state the person is guaranteed to reach the final, no prediction is needed. As the first order development, we might include in the scheme some intermediate state such that transitions from the initial state to this intermediate and from the intermediate to the final are both sure. Then the problem is reduced to finding this intermediate state. Only one such state might be there, because the existence of even one more would provide uncertainty as to which one to choose, so making the prediction incomplete.

Giving the number 0 to the initial state and 1 to the final, let us give $\frac{1}{2}$, say, to this intermediate (no metric is implied – just the order). In the same way we define next $\frac{1}{4}$ and $\frac{3}{4}$ states and so on. This procedure involves only rational numbers, so some infinite sequences of the states might not converge to a state with the rational number to become the initial for the further part of the sequence. Therefore all sequences, i.e., all real numbers are required for guaranteed predictions (Dedekind). In so doing, only order is important, and a state might correspond either to the rational or irrational number as well. Again, no state not belonging to this sequence can exist in the prediction of the steady transition from 0 to 1, otherwise the prediction becomes incomplete. In the Lagrange's version of mechanics, its basic least action principle reflects just this singleness.

Such state sequences are called trajectories, and we are ready now to approach the Newton's scheme, starting with the very condition of the universal predictability. It should be stressed that the scheme is only the necessary language for making universal predictions; it is supported by, though not coming from, our senses that connect us with Nature also in great many other respects.

2 Principles of the Newtonian mechanics

In this essay, I don't consider the post-Newton development of his ideas; even the contribution of Maxwell and Einstein will not be discussed here. My purpose is to understand whether or not the very scheme of mechanics elaborated by Newton is the only possible one. Upon working over many decades in experimental physics, I couldn't refrain from asking myself as to what if there is some other and more efficacious way to address Nature. To this end, I'm going to scrutinize the Newton's scheme in every respect.

Following the method of Descartes of representing geometrical figures with numbers and related equations, Newton has formulated his three 'Laws of Mechanics' in order to apply the similar procedure to physics, i.e., to describe also motion by means of Cartesian coordinates.

The first Newton's law introduces rectilinear and uniform trajectories as free from an external influence ("force"). However, this law is just a vicious circle. As Einstein mentioned in his "The Meaning of Relativity": "The weakness of the principle of inertia lies in this, that it involves an argument in a circle: a mass moves without acceleration if it is sufficiently far from other bodies; we know that it is sufficiently far from other bodies only by the fact that it moves without acceleration."

Aiming at numbering arbitrary motions, we have first of all to match abstract geometric images with real operations. Indeed, what does it mean "rectilinear" in Nature? How rectilinear a trajectory should be for the scheme still being suitable? How to make it sure that a line is straight? Suffices it to be described by linear equations in a reference frame formed as the Cartesian structure? But then, we have to recognize first that our reference frame itself is comprised of straight axes. The commonly accepted agreement suggests using some standard rigid rods. How rigid? Sometimes rigid might appear soft. This depends on the inter-atomic distances, but the concept of distance is still to be introduced using standard rods. (Circle!) We are to transport the rod as along the reference frame axes for marking them evenly, so also over the whole space with parallel shifts and rotations, being sure that it remains rigid. In so doing, we believe that

no actions destroy these operations. The marks on the axes define Cartesian coordinates, which will further be used to define a scalar – squared "length" as the sum of the squared coordinate differences. Only then can we construct the full Cartesian structure using equal length rods to obtain the necessary symmetries of reference frames. (One more circle!) Also collimated light rays might be used, whenever diffraction (still depending on wavelength!) could be neglected, either solely to define linearity, or together with rods for parallelism and other symmetries. Being applied to measurements of motions, we inquire the relevance of these devices, since in fact this procedure has nothing to do with the motion in question. It might well happen that in the study of motion our artificial rods either add something of their own or hide something, so being suitable within only some limited scope of motions. We cannot refer here to great many successful technical applications as well as to the broad experimental support, since all these are carried out within the introduced in advance basic conceptions, so being relevant only within some narrow areas of the implied research.

Even more difficult questions spring up upon considering the time intervals measurement and its universal applicability to real motion. How do we know that the duration of one hour now is equal to that in the future (see, e.g., H.Weyl, "Space-Time-Matter")? How uniform free motion is to be for the scheme to remain suitable? Beginning with Zeno, Aristotle etc., philosophers were burdened by the mystery of time, and Newton himself attempted, in vain, to develop the concept of "genuine" time, that runs uniformly and is free from any influence, our astronomic time being only an approximation of. The summary of his meditations might be found in his "Mathematical Principles of Natural Philosophy": "I do not define time, space, place, and motion, as being well-known to all."

Not belonging to these "all", I want to examine the very necessity of the conventional definitions. Intrinsic to our mind (i.e., being a priory, as in Kant's works) ideas of "space" and "time" suggest only some freedom of motion. However in the Newtonian scheme, the space is already supplied with the three-dimensional Euclidean geometry, that is, it is a somehow defined set of elements – positions – that form the non-compact metric space with all the related properties. The time is not merely "past-now-future" but also a one-dimensional metric space with the countable base of open sets (neighborhoods), and its metrics is monotonous. Why all these?

Imagine a body placed into empty space. How can we tell between its being at rest and moving? The question is quite senseless provided nothing else is there. A reference frame is this "else" in the Newtonian mechanics. Only then can we define the trajectory of this body using readings on the reference frame axes. Still, this frame is only an auxiliary means in the problem. But why do we need to know this trajectory? This becomes meaningful only if some other bodies

may come into contact with this one, and it is this contact that is in question of any real problem in mechanics and generally – in physics.

The purpose actually consists in predictions of the contacts, implying the further action to influence the reaching of this contact. Then, why do we need an intermediary like an external reference frame, rather than to directly consider only the motion of the bodies of interest in our problem? If the event of contact in question does occur, the coordinates of the bodies coincide at some time moment. Hence, the trajectories must (in the Newtonian mechanics) be written in numbers as time-functions of the coordinates taken from the reference frame. Only if times for different trajectories are appropriately coordinated, the predictions of contacts become possible. The accepted solution is one time for all trajectories in the problem, and the synchronized clocks are needed at each position in the reference frame.

All this rather complex measurement system is feasible, provided:

(i) Synchronizing signals connect all positions of the reference frame instantly. Believing that "for any fast motion a faster one might be found", an overcoming signal must always be used, so that observation of the body that could come into contact of interest would never have been lost.

(ii) Suitable clocks are to be made somehow. In daily life rough astronomical timing: years, months, days, hours, might be inappropriate. However, the design of mechanical clocks is based on the previously established principles of mechanics that are still under examination in our essay (One more circle!).

(iii) Identity of the clocks periods is perfect.

The second Newton's law describes some external influence on the trajectory – a force. The idea consists in integrating the series of free trajectories' segments to approximate the actual trajectory as altered from the free motion by this (smooth) force. The end points of each segment contact those of its neighbors. With the reference frame readings their lengths can be used to obtain the measure for integration. The transitions between the segments normalized to the related time intervals define the proportional to the force 'acceleration' as the measure for the transitions between the segments. Leaving aside the mathematical details of these approximations and their limits to the Calculus, I want to focus on the very measurement of a force in Newtonian mechanics. Indeed, where to find the vector of the force? Traditionally, some particular kind of forces is suggested for the problem of interest like the gradient of an external potential (as, e.g., in oscillations, gravity), friction, electromagnetic field etc. There is no general concept of force in the geometrical terms of the scheme itself. Provided the force is given in advance all over space-time, the whole trajectory can be found step by step. However, this approach cannot produce a genuine prediction as yet, being dependent on the knowledge of force up to the final state where no prediction is already in-

teresting. In the Newtonian mechanics, inertia determining acceleration makes the scheme really predictable: Given the force, a sufficiently big mass of the body will send this force to the second order perturbation in the trajectory determining equation. It is just the demand of predictability that is responsible for second order terms in the equations to be sufficient: Force collected over the first order linear segment provides the next inter-segment transition, and no higher order terms are needed to determine them. So, the specification of only the initial free segment suffices to predict the final contact. This fact is not always understood, especially by mathematicians, believing in the known from experiment harmony of Nature. For instance, V. Arnold in his famous textbook "Mathematical Methods of Classical Mechanics" declared: "It is possible to fancy a world, in which for the determination of the future of a system one has to know in the initial moment also acceleration. Experiment shows that our world is not such."

However, any statement and result of experiment is formulated in terms of the already accepted theoretical models (Einstein: "In order to measure the velocity of light, the theoretical concept of velocity is necessary."). All these concepts originate in predictability. As a matter of fact, there is no harmony in our world, but the demand for predictability bounds us to develop a scheme ready for advising the person, looking obliviously around for the solution of his problem, to try first of all physics for the reaching of his wanted state.

The third Newton's law introduces the concept of interaction between bodies as a sole source of force, so providing some certainty to the second Newton's law. Then, an isolated from external influences collection of either or not interacting bodies taken as a whole must move freely according to the first Newtonian law. In particular, a solid body, considered as comprised of two parts separated with an infinitesimally thin gap, moves freely while, according to the second Newton's law, an additional force would be needed to keep each part moving free in spite of their reciprocal attraction. Hence we have to admit that the action of one part on the other is compensated by the opposite action.

3 Alternative numbering of motion

Newton considered velocities of bodies extendable in their values up to infinity, and then the using of located in advance clocks and rods became indispensible. Success in geometry tempted the using of the trajectory as the basic entity to start a theory with. On the contrary, the existence of the top-speed signal makes it possible to suggest a different numbering of motion. In so doing, we need no metric – no rods, no clocks, no material points, no reference frames. Our main concept is "contact", defined solely by its existence – "yes/no". The concept of body will be used just as a picturesque representation of contacts. It is the prediction of a contact using some auxiliary contacts – the Contact Problem (CP), that is the only

issue of physics as a method to make universal predictions whenever relevant.

Attempts to define the space-time geometry with trajectories of limited velocities have been carried out in the middle of the past century [1-6]. In the interior of the light cone, trajectories were used to define neighborhoods generating the space-time topology as sets of points (events) such that any trajectory reaching a point of the neighborhood starting from outside passes also some other points of it, and there is some open interval in the order of the 1-dimensional continuum of this trajectory contained in this neighborhood (see Ref. 7 for details).

Consider two bodies A and B moving, each one along its (ordered) trajectory, toward their possible contact denoted (A,B). Let a set of auxiliary bodies be simultaneously emitted from A so that some of them reach B. Find the first of them to come into contact with B in the own B-order (One might imagine this first to put a mark on B, so that others meet B already marked.). Such a body will be taken for the top-speed signal, provided the emitted set is rich enough to cover all possible applications. A top speed must exist in the scheme for B not to be lost from observation upon its accelerations, so making predictions impossible. In so doing, we don't provide this top speed with a numerical value (no cm/sec, just topmost as defined!). Let further B emit instantly in response a similar set to reach A; it might be regarded as 'reflected' from B. This procedure being multiple repeated will be called the oscillation of the top-speed signal between A and B.

Our scheme of numbering motion consists solely in counting the numbers of these oscillations n_{AB}. Let us start this counting at some state of A. If (A,B) exists, the number of the oscillations is infinite, since were it finite some last oscillation before (A,B) will be there, in contradiction with the top-speed property of the signal, since either A or B would then reach (A,B) sooner. It is tempting to take the infinity of nAB for the prediction of the contact, but in the absence of (A,B) this number is still infinite though in the Newtonian scheme it would take infinite time; but we claim to use no measure for time, only the order.

In order to obtain the prediction, we can use an auxiliary body X with (A,X) known in advance and measure the ratio n_{AB}/n_{AC} for the triple (A,B,X), beginning at arbitrary point. (Both numbers being infinite, the ratio doesn't depend on this point.) The prediction of (A,B) follows from that of (A,B,X) provided such X can be found that this ratio is finite. Again, this is not a genuine prediction as yet, because we are counting the ratio up to the (A,B,X), and then nothing is left to predict. Hence, a scheme is to be developed to predict (A,B) already at the beginning of the oscillation numbers (ON) counting. Although we dispensed with all Newtonian intermediaries and turned to measure a motion solely by means of some auxiliary motions, we have yet to develop a scheme similar to the Newtonian to obtain genuine predictability.

For this to be possible, we ought now to consider suitable for our numbering scheme intersections of trajectories that allow for using the related concept of force. To this end, we define first the class Q of trajectories, the contacts between which are not too dense, so that with ON counting it be always possible to distinguish contacts however multiple. For instance, two trajectories, which in the Newtonian version have contacts only in all points with rational values of even one of coordinates, don't belong to Q. Hence, if trajectories from Q have two or more mutual contacts, ON counting, wherever started, might become infinite for only one of these. Only trajectories from Q are suitable for CP.

If the top-speed body signal S emitted from A at some of its point to contact B at some of its point, then no body emitted from A simultaneously with S can contact B in all points earlier than (S,B) in the B-order. So, we have now points in A and B that cannot be connected with trajectories unlike that in the Newtonian scheme. The set of all points, no pair of which can be so connected is called "spacelike hypersurface" W, and its elements will be called positions; therefore the trajectory of A, say, can contact W only at a single position. In particular, all top-speed signals connecting a point of A apart from W define some boundary in W: Only positions of W within this boundary can be connected with the part of A bounded by this point. An open in its order interval of A, crossing W at some of its points can be projected on W inside this boundary. This can be done using a series of mutually "parallel" trajectories (The notion of parallelism might be defined using a system of four ratios of ON's, and so defined parallel trajectories are not necessarily straight lines.) as follows. Take r points on A such that the finite ON's between neighboring pairs of parallel trajectories, connecting them to W, differ by only one oscillation. Increase r keeping this condition. In the limit r going to infinity we obtain a path of positions in W, which are in one-to-one correspondence with the set of r points in A trajectory to form (again being completed with irrational limits) the one-dimensional continuum. Unlike trajectories, paths might have self-intersections, though "rarefied" in accord with the trajectories they are projections of.

The whole W is an "envelope" for various combinations of possibly intersecting paths. If paths intersect, then the contact of their trajectories either exists or not. However, if paths don't intersect no contact can be there. It is only this purely topological property that is important for CP. W must have enough freedom to allow all the variety of combinations of passes. Since paths and their allowed combinations are one-dimensional, they might be topologically embedded in the 3-dimensional Euclidean space (Remember traffic interchanges. In general, a wide class of n-dimensional spaces, including our paths, might be so embedded in the Euclidean space of the dimension 2n+1, according to the Noebeling-Pontryagin theorem. Hence the geometry of space, taken in the Newtonian scheme as fallen from heaven, merely results

from the union of all paths, and more dimensions for W would be redundant, because already some 3-dimensional subspace of it can include all cases for CP. Importantly, W cannot be considered as a sub-space of the 4-dimentional Lorentz spacetime, otherwise its meaningless topology with non-countable neighborhoods would be only 1-dimensional in both Lebesgue and Poincare senses.

A top-speed signal cannot have more than one contact with any other trajectory in our scheme. Some other trajectories might have single contacts too, and these will be useful to define a force. Let us therefore select a special class of trajectories – the measurement X-kit with the following properties:

(i) Two trajectories from X either have no contacts or have only one;

(ii) Any point of a trajectory from Q has contacts with some trajectories from X.

(iii) Any two points of a trajectory from Q can be connected by a trajectory from X. Free trajectories of the first Newton's law are such, and just these properties of them, perhaps only locally, are actually needed in our scheme too.

In the second Newton's law acceleration is determined by force. Let us now inverse this law so as to determine force via acceleration, though not of the body of interest in the CP but of a body from the specially prepared auxiliary test P-kit with the same scheme of contacts as the X-kit, however comprised of bodies with some fixed constants to be specified for the particular kind of forces. Provided such standard constants exist over the whole Q, one is able to determine the acceleration of the body A that is of interest in CP comparing its acceleration at each point to that of the test body from the P-kit here, given the related constants of both. If the bodies participating in this comparison differ from each other only by the values of their constants, the trajectory of A can be defined, and therefore it is worthwhile to represent a force as the product of a constant and an entity defined by the ON counting – field. With the definition of our two kits, the said comparison might always be achieved with the counting of ON's and their ratios. The mentioned properties of the kits are specified just to allow for this comparison, so defining situations, in which we claim to make reliable predictions.

In the chain of links approximating a trajectory with a broken line, it is sufficient to specify only the first link. Then the force defining inter-link transitions (given the required constants) provide the prediction.

It remains now to define the required constants in terms of ON. We specify first a regular P-star, comprised of trajectories of some P's from the P-kit with the common contact, in which the ON ratios are distributed regularly:

(i) Each trajectory of P has the neighbors, that is, a number of trajectories, the ratios of the ON between P and any its neighbor to that between P and any other trajectory from the star exceeds 1; it follows that the ratios of the ON between P and any pair of its neighbors equals 1.

(ii) This feature is the same for all trajectories of the star.

In 3-dimensional Q these conditions can be exact only for a star with the configuration of a Platonic solid (If a star comprises great many trajectories, this inexactness might be ignored in the definition of a measure as the numbers of trajectories in subsets of the star; this is used, e.g., in the problems of field propagation, however not referred to further on in this essay restricted to mechanics.).

Consider a Platonic solid star with the bodies from the P-kit moving from its vertices toward the center solely under their interactions (Remember the third Newton law.). It is convenient to describe the gauge procedure for the constants in Newtonian terms (translation into the ON counting will be evident). These bodies are assumed to have some masses m and charges q. The completely identical bodies can reach the center only being mutually attracted as for gravity; otherwise some charge compensation is needed. Then only two of the Platonic solids might be relevant: the cube and the icosahedron. Indeed, in both it is possible to distribute opposite charges so as to obtain a regular star for bodies from P-kit.

The cube might be arranged out of two interwoven tetrahedrons – one with $+q$, another with $−q$; hence the star is neutral as a whole. All 8 initial velocities are radial and equal, and 8 equal initial radii are also the same for all bodies of the cube. All these bodies are being equally accelerated proportionally to q^2/m toward the center along rays, whatever radial dependence of their (isotropic) interaction force. We ascribe the cube star to electromagnetic (EM) interaction, the magnetic component of which is then equal 0 on the rays, and the electric field is purely radial.

Starting ON counting from the initial radii, we find their ratios for each ray with its neighbors to be 1 for any n. Reversing argument, the value 1 of these ratios can be taken as the criterion for the cube star to be perfect. After passing the star center the bodies decelerate to reach initial velocity at the same radii as the initial ones. Here some of them can be used, with an appropriate order of the vertices, to form the descendant star from this seed, adding more similar bodies. A triple of the neighboring seed star bodies completely determines all other members of the descending star with ON ratios counting. In the progress of this descending step by step in all directions, the charge and mass are transported over the whole network in Q, so determining the same pair of standard constants everywhere. Importantly, both m and q must be the same in the cube: Varying any of them in a part of cube, even keeping the value q^2/m unchanged, destroys the star symmetry. Hence, the network transports both standards unchanged.

In a more general case of CP, e.g., with an arbitrary external EM field, the source of which is not known in advance, unlike that in the Newtonian approach, the acceleration of charged bodies is proportional to the q/m rather than to q^2/m. However, the value of q/m is also determined by the cube star gauge, since both q and m are preserved upon the descent transportation. So, predictions based on ON counting are available in CP even beyond the Newtonian scheme.

The icosahedron regular star of oppositely charged bodies (also neutral as a whole) exists only if, in the Newtonian sense, the interaction force increases with radius. Whereas the cube is a sub-star of the full dodecahedron, the icosahedron stands alone; hence its charge and mass have nothing in common with EM q and m. With the distance increasing of its force, allowing for confinement and asymptotic freedom, the icosahedron star symmetry might be suggested to explain the Dark Cold Matter and the Dark Energy in cosmology.

4 Postscript

The origin of the "Laws of Nature" for any method of numbering motion as well as of the concept of motion itself results merely from the very problem statement by the person-user to find, whenever possible, a universally predictable course of action. To this end, physics suggests CP. Nature has no harmony of its own; only living creatures are looking for reliable schemes to make predictions. In particular, it is clear now why quantum mechanics had not developed its own variables instead of classical position and momentum. However modified, these variables still present information in terms required by the user.

References

1. Zeeman E.C. *Topology*, 1967, v. 6, 161.

2. Tselnik F. *Sov. Math. Dokl.*, 1968, v. 9, 1151.

3. Hawking S.W., King A.R., McCarty P.J. *J. Math. Phys.*, 1976, v. 17, 174.

4. Gobel R. *Comm. Math. Phys.*, 1976, v. 46, 289.

5. Ehlers J., Pirani F.A.E. and Schild A. In: General relativity, ed. L. O'Raifeartaigh. (1972).

6. Marzke R.F., Wheeler J.A. In: Gravitation and relativity, ed. H.Y. Chiu and W.F. Hoffmann, New York, (1964).

7. Tselnik F. *Progress in Physics*, 2016, v. 12, no. 2.

The Interpretation of Sound on the Basis of the Differentiated Structure of Three-Dimensional Space

Gerhard Dorda

Institute of Physics, University of Armed Forces Muenchen, Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany.
E-mail: physik@unibw.de

It is shown that the experimental data of sound, obtained by the investigation of H. Fletcher and W. A. Munson [4], can be physically described on the basis of the differentiated structure of three-dimensional space (DSS), showing an analogy to the physical interpretation of the process of human vision. The analysis of the experimental data indicates that the process of hearing at frequencies below 800 Hz depends on the differentiated structure of the space related to air. Furthermore, it has been shown that the existence of sound at frequencies higher than 800 Hz is the result of quantization phenomena of the differentiated space-related state of the air, revealing to be an analogy to the quantum effects of the differentiated structure of space of the quantum-Hall-effect (QHE). The presented results about sound, considered with respect to the findings of the QHE, the Hubble-effect galaxies and the process of seeing, result in the fundamental statement that the human ability of the observation of being refers exclusively to the existence of the differentiated structure of three-dimensional space.

1 Introduction

The discovery of a macroscopic quantization in the field of solid state physics, called quantum-Hall-effect (QHE) or Klitzing-effect [1], which was first experimentally observed by K. von Klitzing in 1980, opened the door to a new interpretation of various physical phenomena, such as the origin of the category time or dynamics in the field of mechanics, or thermodynamics and theory of heat [2], but also to some human-related biological processes [3]. The experimental findings of the QHE provided basic indications of the possibility of the existence of a specific space state, characterized by a division of three-dimensional space into a clearly separated, independent 2-D and 1-D dimensional space, called *Differentiated Structure of Space (DSS)* [2, 3]. This separation is recognizable e.g. by the simultaneous existence of two different forms of electromagnetism, effective not only in the context of MOS transistors, but also in the observation of the Hubble-effect (HE) galaxies, a process that is even reflected in the process of human vision, among other things. As shown in [3], the fundamental investigations into the existence of HE galaxies lead to the physical realization that the vision of humans, and to some extent also of animals, depends on the given DSS-state of space. The fundamental importance of the DSS-space state for humans becomes additionally apparent when we discuss the sound process physically. The experimental data of the investigations of human hearing carried out by H. Fletcher & W. A. Munson in 1933 [4], which show the relationships found between the sound pressure, the sound intensity and the loudness level on the one hand and the sound frequency on the other hand, presented here in Fig. 1, have so far been interpreted as biologically caused effects [5–7]. In contrast, the presented work shows that all the dependencies measured by Fletcher & Munson [4] (except for the conditions at the initial and final frequencies of human sound sensitivity are almost exclusively of physical origin, since, as is shown, they are due solely to the existence of the DSS-state of the air atmosphere.

Based on the data in Fig. 1, it must first be pointed out that for the investigation of sound intensity and sound pressure at the boundary condition of approximately 20° C and a sound velocity of v_s = 343 m/s, the frequency of f_0 = 800 Hz proves to be a suitable boundary condition, since the physical processes involved in the realization of sound, arising at frequencies $f_x < f_0$, differ considerably from the processes at frequencies $f_x > f_0$. Therefore, we divide the physical analysis of the sound process into Part I and Part II.

2 The analysis of the Part I area of sound

Sound generation and its transmission are based on the properties of air. The air molecules as components of the air, which we may evaluate as an ideal gas in the closest approximation, are mainly subject to the influence of earth gravity. Since these forces can be regarded as constant in wide areas above the earth's surface, there is the special possibility of not paying attention to the gravitational forces when analysing the origin of sound. Following this idea, we can therefore assume that in our case the kinetic energy of the air molecules and their variability can be considered as purely electromagnetic in nature, which, however, as the experimental data show, is causally related to the temperature of the environment, i.e. more precisely, the fundamental electromagnetic energy of the air molecules is indirectly proportional to the ambient temperature, observable especially in the variable value of the speed of sound. This in turn means that the air can be considered a so-called Boltzmann gas, i.e. the electromagnetic energy of the air molecules can be put into a causal
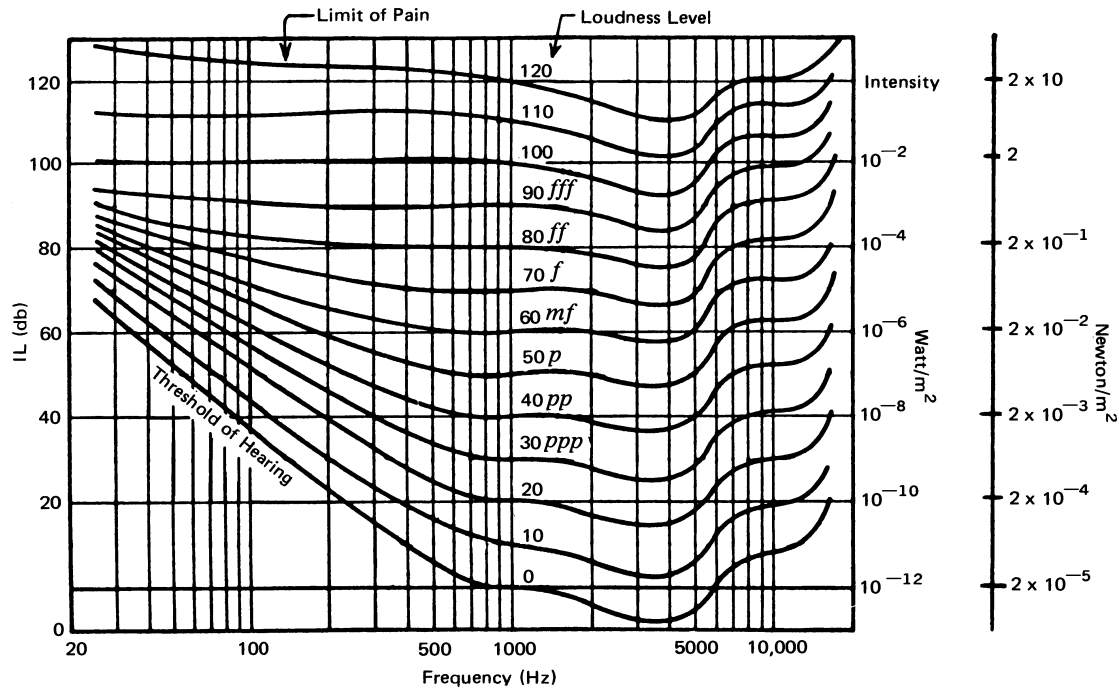
Fig. 1: Curves of equal loudness (Fletcher and Munson [4]). Reprinted from [6] by permission from *Springer Nature* (`https://rd.springer.com/chapter/10.1007/978-1-4615-9981-4_3`). This figure was first published by the *Journal of the Acoustical Society of America* and later adapted and expanded by *Springer Nature*.

relation to the Boltzmann energy kT. Based on these considerations and on the experimental data of the sound investigations of Fletcher & Munson, the following basic equation is formulated for the analysis of sound:

$$\frac{kT}{hf_x} = n_{B,x}\, v_s \frac{p_x}{I_x} = n_{B,x} \frac{E_{p,x}}{E_{I,x}}\,. \tag{1}$$

Here, k is the Boltzmann constant, T the temperature of the environment, h the Planck constant, f_x the given frequency, v_s the speed of sound, p_x the sound pressure, I_x the sound intensity, and $n_{B,x}$ the number of air molecules corresponding to the given frequency in order to reach the so-called hearing threshold, as shown in Fig. 1 [4]. The index x in (1) refers to the given frequency f_x for all quantities. The sound intensity energy $E_{I,x}$ occurring in (1) is given by $E_{I,x} = I_x/f_x$, and the corresponding sound pressure energy by $E_{p,x} = p_x l_x$, taking into account the sound velocity of $l_x f_x = v_s = 343$ m/s, l_x represents the so-called sound length.

As can be seen from the definition of a so-called Boltzmann relation of sound presented in (1), this relation at the frequency f_x is given by the relation of the corresponding energies, which in turn can be determined by the indirect proportionality between the sound pressure p_x and the sound intensity I_x. In order to be in accordance with the experimental findings, the energy relation was, as shown in (1), additionally modified by the number of molecules $n_{B,x}$ considered at the given frequency f_x in order to be able to causally represent the

variation of the energy relation. The fundamental importance and necessity of the introduction of the number $n_{B,x}$ will be presented in the following analysis, because this number $n_{B,x}$ is not only of decisive importance in the description of the course of the (human) hearing threshold at the frequencies of approximately 20 Hz $< f_x < 800$ Hz, but it is the essential factor that helps to physically fathom the process of sound realization in nature.

In the following it is shown that the special form of the Boltzmann relation equation formulated in (1) can and must be used as a starting point for sound analysis.

3 The relationship between sound and the air-related DSS-condition

The experimental results of Fletcher & Munson [4] are not only fundamental for the description of human hearing, but, as is shown, are generally valid and therefore fundamental [5–7]. In the analysis to interpret the realization and propagation of sound, the following experimentally observed facts must be considered:

1) The sound intensity I_x and the sound pressure p_x are in a mutually causal relationship at any volume (indicated in phon*) and at any observed frequency. Special attention must be paid to the finding which reveals that a constant connection

---

*The phon is a unit of perceived loudness of pure tones, indicated in Fig. 1 as "Loudness Level".

is given between the square of the sound pressure, i.e. p_x^2, and the sound intensity I_x. This physically conditioned circumstance, which is valid for each sound intensity I_x and for each corresponding sound pressure p_x, namely at each frequency f_x [4], see Fig. 1, can be described by

$$\frac{p_x^2}{I_x} = \text{const.} \,. \tag{2}$$

This means that with an increase in sound intensity I_x (phon strength), the increase in sound pressure p_x must always be smaller than the increase in the associated sound intensity I_x.

2) It should also be noted that the curves of equal loudness dependent on f_x, shown in Fig. 1 in phon, do not touch each other, i.e. the intensity distances, independent of the frequency(!), have almost the same values, i.e. the increase of the I_x-phon distances is almost always constant at a given frequency f_x. This important experimental finding indicates that the so-called loudness of the sound refers solely to the sound intensity I_x.

These two experimental findings show that the energy values characteristic of the air molecules, expressed in purely electromagnetic form using $E_{I,x} = I_x / F_x$ and $E_{p,x} = p_x l_x$, can be recorded and expressed in two different ways when sound is observed.

Here are a few remarks: The specific type of the DSS-air-state (Differentiated Space Structure state of air), seen in terms of the boundary conditions of the existence of the Hubble-effect (HE) galaxies [3], is to be sought in the theoretically possible cancellation of the gravitational effect of the Earth as a boundary condition for the emergence of the specific type of the DSS-air-state, described in the beginning of Part I.

The essential consequence of this cancellation is that the variability of the sound-related energies of the air molecules is limited to changes of electromagnetic nature alone. This means that, in contrast to the states observed in the HE galaxies, the molecules of air in the given DSS-state in this case show only purely electromagnetic variable effects, and that within the framework of our common three-dimensional understanding of space, they are separated into a so-called 2-D and a 1-D "space".

The special feature of this insight is that it clearly reveals for the first time that, "spatially" considered, air molecules as energy carriers can be observed in two different forms, i.e. "spatially" differentiated, whereas in our common, i.e. classical understanding of space, on the one hand the $E_{I,x}$-energy refers to the two-dimensionality of this energy, i.e. to the 2-D "space", which is interpreted as intensity energy, and on the other hand the $E_{p,x}$-energy refers to the one-dimensionality, i.e. to the 1-D "space", which is interpreted as pressure energy. However, the possibility of the existence of such a special state is only given if we may consider nature, evaluating in spatial categories, as differentiated, recorded as DSS-state.

This specific spatial state was first discussed in 2017 in [2, pp. 33–34, 45, and 49–50], based on the analysis of the experimental findings of the quantum-Hall-effect published in 1980 by K. von Klitzing et al. [1], and its unusual existence was again proven by the analysis of HE galaxies [3]. Seen in this context, the analysis of the Fletcher & Munson data shown in Fig. 1 and its conclusion are of extraordinary importance, because they show that the possibility of sound formation is only given when this specific "differentiated space state", unusual to our daily understanding of space, is given in the air and thus the boundary condition for the formation of sound is real. The correctness of such an unusual model, which was presented on the basis of the specific electromagnetic DSS-state of the air, can be confirmed impressively and convincingly by a further detailed analysis of the experimental data of Fletcher & Munson.

The essential functional significance of the number of molecules $n_{B,x}$ given in (1) at the given intensity I_x is to guarantee the DSS-state of the air in the form of (1) and (2), which physically reflect the limit value of the hearing threshold as equations. This in turn means that, for the experimentally given values of sound intensity I_x and sound pressure p_x, we can use (1) to write the specific magnitude of the value $n_{B,x}$ as a function of the frequency f_x, where we have to define the frequency f_x, normalized in relation to the reference frequency f_0, as the relation $n_{f,x} = f_x / f_0$. And in order to be able to mathematically record the homogeneity of the air in the DSS-state, it is also necessary to define the relation $n_{p,x} = p_x / p_0$, i.e. to set the sound pressure p_x in relation to the standard value of the air pressure p_0 by means of the number $n_{p,x}$. Starting from the limit values $I_0 = 1 \times 10^{-12}\,\text{W/m}^2$ and $p_0 = 2 \times 10^{-5}\,\text{N/m}^2$ given in Fig. 1, for every frequency we then obtain the constant value $(n_{B,x}/n_{f,x}\, n_{p,x}) = 1.122$. The constancy of this value, which captures the homogeneity of the air condition, and above all the small size of this dimensionless value 1.122 suggests the possibility of replacing this value by the number 1, followed by the associated necessity to modify the experimentally given values I_x and p_x accordingly. In order to minimize the change of the value p_x^2/I_x, the validity of which must be maintained, it is sufficient to reduce the value from $I_0 = 1 \times 10^{-12}\,\text{W/m}^2$ to $I_0 = 0.9 \times 10^{-12}\,\text{W/m}^2$, while keeping the value $p_0 = 2 \times 10^{-5}\,\text{N/m}^2$. The smallness of the correction of the I_0 value is fully acceptable, as it is within the given measuring accuracy.

If we now try to represent a causal connection of the values $n_{B,x}/n_{f,x}$ to the experimental values $n_{\exp,x} = I_x/I_0$, which are given by the known data of [4], by means of an equation, an extraordinarily meaningful connection, valid for all frequencies $f_x < f_0$ emerges, which can be described by

$$n_{\exp,x} = \left(\frac{n_{B,x}}{n_{f,x}}\right)^2 . \tag{3}$$

Using a simple calculation, it can be shown that (3) is a com-

pellingly necessary causal consequence of the hearing process (1) and (2).

Since in the framework of the DSS-model, the values of $n_{\text{exp},x}$ readable in Fig. 1 can be unambiguously related to the specific energies of the 2-D space, i.e. to the sound-intensity radiations I_x, but the relation values $n_{\text{B},x}/n_{f,x}$ within the framework of $(n_{\text{B},x}/n_{f,x}\,n_{p,x}) = 1$ can be related to the specific energies of the 1-D space, i.e. with the sound-pressure values p_x, the simple and clear form of (3) proves that every frequency $x < 800$ Hz must be the condition of the differentiated state of space. Furthermore, (3) testifies that this DSS-state is necessary as a boundary condition in order to reach the sound limit by means of the I_x and p_x limit values, i.e. to generate and transmit sound in our environment of the earth's surface.

The number $n_{\text{B},x}$ in (1) is also of great importance for non-physical reasons: As shown, it is necessary for the realization of the DSS-state and thus brings the mental development of man to fruition. The analysis of the sound indirectly shows clearly that the acoustic communication between humans is solely caused by the existence of this DSS-state of the air. In fact, an interesting analogy to the process of seeing and thus to the human perception process in general can be seen, because the process of seeing, as shown in [3] and explained by the existence of specifically suitable uvula and rod cells, is also based on the existence of the DSS-state, as it were, in the field of optics, i.e. light.

4 The analysis of the Part II region of sound and the analogies to the integral and fractional quantum-Hall-effect

On the basis of the sound interpretation model presented in Part I, it is clear that sound mediation at frequencies f_x above the limit value f_0, i.e. at $f_x > f_0$, must be fundamentally different from the process presented in Part I, because once the sound intensity value I_0 is reached, there should be no further normal possibility of reaching the DSS-state for the production of sound. In fact, however, it is observed that initially, with increasing values from f_0 to approximately $f_x = 1300$ Hz, the hearing threshold limit value of I_0, given by approximately 1×10^{-12} W/m², remains quasi-constant in order to reach a new minimum in I_x of approximately 2×10^{-13} W/m² at $f_x = 3200$ Hz. Afterwards, starting at $f_x = 3200$ Hz, an increase in the hearing threshold limit values I_x is observed with increasing f_x values, followed – which is particularly important – by an indication of a small decrease in the I_x limit values at $f_x = 1.28 \times 10^4$ Hz. After that, a strong increase of the curves of equal loudness is measured further on with increasing frequency, and to stop at $f_x = 1.6 \times 10^4$ Hz, at all phon levels. In order to be able to interpret these experimentally observed complex I_x-f_x-dependencies, we have to assume the existence of two different processes which, as we will show, relate to the 2-D space component on the one hand and to the 1-D space component on the other.

One process concerns the interpretation of the minima of I_x at $f_x = 3.2x10^3$ Hz and at $f_x = 1.28 \times 10^4$ Hz: They can be interpreted as the consequence of an area quantization given in the two-dimensional space part, describable with the quantum number 4 and 16. With this, it is postulated here that – despite the sound limit value of $I_0 = 1 \times 10^{-12}$ W, as explained in Part I – also at higher frequencies, i.e. at $f_x = 3.2 \times 10^3$ Hz and $f_x = 1.28 \times 10^4$ Hz – which is due to this experimentally observable macroscopic 2-D quantization – an air-related DSS-state can be present, which means that in nature it is possible to also generate sound at $f_x > 8 \times 10^2$ Hz.

An analogous macroscopic quantization related to two-dimensional space, namely the quantization discovered by K. von Klitzing, was renamed in later years to *Integral quantum-Hall-effect* (IQHE), observed at the quantum numbers 2, 4, (6), 8, 12 and 16 [2, 8, 9]. In order to be able to consider the assumed two-dimensional surface quantization for sound as physically acceptable in comparison with IQHE quantization, some additional remarks are necessary: In the IQHE, the magnetic field B is in causal interaction with the electron density N_{e}, i.e. in the DSS-space model with the 2-D space state. In the so-called Fractional quantum-Hall-effect (FQHE), discovered for the first time by D. C. Tsui et al. at GaAs–Al$_x$Ga$_{(1-x)}$As heterostructures [10], the magnetic field B_x corresponds to a frequency f_x. This model of the different functioning of the magnetic field, given on the one hand by the IQHE and on the other hand by the FQHE, could actually be indirectly confirmed by targeted measurements within the QHE, as shown in [11, pp. 34–42]. This means that the magnitude of the magnetic field, which is expressed in Tesla units in the MKSA unit system, can also be expressed simply by the quantity "frequency" in the MKS unit system, which makes a possible analogy of the process between the sound effect and the QHE appear possible.

As already mentioned in [3], the QHE state is always present in the DSS-space state. As a consequence, IQHE quantization is to be interpreted as a 2-D space quantization, in contrast to FQHE quantization, which can be interpreted as a 1-D space quantization. This insight leads us to the additional conclusion that the discovery of the length-related harmony theory, which stems from Pythagoras, actually reflects a 1-D space quantization, which is today presented in every musical harmony theory as a consequence of the existence of overtones that always belong to the fundamental tones. But this important insight must be further expanded by the discovery of the existence of deep harmony tones associated with each fundamental tone, recognizable by the existence of the so-called deep combination tones, see [6, p. 38]. It is evident that the existence of these deep harmony tones can be understood as an analogy to the existence of FQHE quantization. This leads to the conclusion, which is important for our analysis, that the unexpected sound generation at frequencies $f_x > 8 \times 10^2$ Hz according to our model must be a consequence of the existence of 2-D space and 1-D space quanti-

zations associated with the given fundamental tones, a model which is fully consistent with the extensive experimental findings of both the IQHE and the FQHE. In addition, an interesting fact can be seen that the QHE, where the observation of the FQHE was initially completely unexpected, fully reflects this "unexpectedness" when listening to deep harmony tones [6]. The relatively small probability of low combination tones therefore means that a strong increase in the hearing threshold limit values I_x is to be expected with increasing frequency, which, as Fig. 1 shows, was actually observed.

The found sudden stop of sound generation at $f_x = 1.6 \times 10^4$ Hz at all phon values can only be interpreted in such a way that in humans in the cochlea there are no stereocilia for these high frequencies that would process such electromagnetic signals. This means that we can only speak of a physiological effect in this case of the general cessation of sound sensitivity at extremely high frequencies. The same reasoning can be applied to the description of the sudden occurrence of hearing ability observed at all phon levels, which occurs in humans at about 20 Hz. This means that the onset of hearing must be physiological and therefore cannot be attributed to a physical effect. Otherwise, as explained, the sound data of Fig. 1 observed by Fletcher & Munson can be attributed to physical processes, which all, without exception, indicate the existence of an air-related DSS-condition.

5 Summary and conclusion

Based on the analysis of experimental data of the quantum-Hall-effect [1], it was found that in nature, spatially speaking, a specific state can exist, called differentiated structure of three-dimensional space (DSS-state) [2, 3]. Based on this discovery not only a novel description of the category "time" as a consequence of localized, i.e. 1-D related electromagnetism could be presented [2, page 45], but also the background of the existence of the Hubble-effect galaxies as well as the process of human vision based on the DSS-state could be physically described [3]. This visual model, which associates the rod cells with the specific 1-D space state and the uvula cells with the specific 2-D space state, does not differ in any essential point from the process of human hearing based on the process of the DSS-state, as the analysis of the Fletcher & Munson data reveal, revealing also the analogy between the processes of sound generation and those of the quantum Hall effect. Thus, within the DSS-model, the air molecules are the "carriers" of both the 1-D space structures in terms of sound pressure p_x, and the 2-D space structures in terms of sound intensity I_x. The detailed analysis of the Fletcher & Munson data also clearly indicated that the limit of the hearing threshold is determined by the existence of the DSS-air-state. This experimental discovery is a fundamentally important discovery from a physical point of view because it proves that the process of hearing is conditioned by the existence of the DSS-state of the air molecules. Conse-

quently, it can be concluded that the DSS-state as the basis of hearing, but also of seeing, as shown in [3], is the fundamental state that enables human kind to mentally recognize what is happening in nature, i.e. all being. But this also means that the DSS-state is the fundamental physical background which is the starting point for all human evaluations and interpretations of both static and dynamic, i.e. time-related processes in nature [2, 3, 11] and should therefore always be taken into consideration additionally.

References

1. von Klitzing K., Dorda G. and Pepper M. New Method for high-accuracy determination of fine-structure constant based on quantized Hall resistance. Phys. Rev. Lett., 1980, v. 45, 494–497.

2. Dorda G. Die Struktur von Raum und Zeit und die Interpretation der Wärme. Cuvillier Verlag, Goettingen, ISBN 978-3-7369-9388-4, eISBN 978-3-7369-8388-5, 2016.

3. Dorda G. The Interpretation of the Hubble-Effect and of Human Vision Based on the Differentiated Structure of Space. Prog. Phys., 2020, v. 16 (1), 3–9.

4. Fletcher H. and Munson W. A. Loudness, its definition, measurement and calculation. J. Acoust. Soc. Am., 1933, v. 5, 82.

5. Mende D. and Simon G. Physik. Gleichungen und Tabellen. Carl Hanser Verlag, Fachbuchverlag Leipzig, Muenchen, 2013.

6. Roederer J. G. Introduction to the Physics and Phsychophysics of Music. Heidelberg Science Library, Springer-Verlag, New York, 1973. In german: Physikalische und psychoakustische Grundlagen der Musik. Springer-Verlag, Berlin, Heidelberg, 1977.

7. Breuer H. and Breuer R. Atlas zur Physik. Deutscher Taschenbuch Verlag GmbH & Co., Muenchen, 1987/1988.

8. Wittmann F. Magnetotransport am zweidimensionalen Elektronensystem von Silizium-MOS-Inversionsschichten. Dissertation, University of Armed Forces, Muenchen, Institute of Physics, 1992.

9. Suen Y. W., Jo J., Santos M. B., Engel L. W., Hwang S. W. and Shayegan M. Missing integral quantum Hall effect in a wide single quantum well. Phys. Rev. B, 1991, v. 44, 5947–5950.

10. Tsui W. A., Störmer H. L. and Gossard A. C. Two-Dimensional Magnetotransport in the Extreme Quantum Limit. Phys. Rev. Lett., 1982, v. 48, 1559–1561.

11. Dorda G. Quantisierte Zeit und die Vereinheitlichung von Gravitation und Elektromagnetismus. Cuvillier Verlag, Goettingen, ISBN 978-3-86955-240-8, 2010. In particular: pp. 34–42, Fig. 2.1 and 2.2, and Eq. (2.14).

A Pedestrian Derivation of
Heisenberg's Uncertainty Principle on
Stochastic Phase-Space

G. G. Nyambuya

National University of Science and Technology, Faculty of Applied Sciences – Department of Applied Physics,
Fundamental Theoretical and Astrophysics Group, P. O. Box 939, Ascot, Bulawayo, Republic of Zimbabwe.
E-mail: physicist.ggn@gmail.com

Without using the common methodologies of quantum mechanics – albeit, methodologies that always involve some demanding mathematical concepts, we herein demonstrate that one can derive in a very natural, logical and trivial manner, Heisenberg's quantum mechanical uncertainty principle on the new phase-space whose name we have herein coined *Stochastic Phase-Space*. This stochastic phase-space – is a mathematical space upon which we previously demonstrated [2] the naturally implied existence of the First Law of Thermodynamics from Liouville's theorem. In addition to Heisenberg's uncertainty principle, we derive an upper limiting uncertainty principle and it is seen that this upper limiting uncertainty principle describes non-ponderable tachyonic particles.

It must have been one evening after midnight when I suddenly remembered my conversation with Einstein and particularly his statement, 'It is the theory which decides what we can observe.' I was immediately convinced that the key to the gate that had been closed for so long must be sought right here. I decided to go on a nocturnal walk through Faelled Park and to think further about the matter …Werner Karl Heisenberg (1901-1976). Adapted from [3, p. 6].

1 Introduction

The present paper is the third in a five part series where we make the endeavour to understand the meaning and origins of what drives the unidirectional forward arrow of thermodynamic entropy. In our first instalment [4, hereafter Paper I], we demonstrated that the *Second Law of Thermodynamics* (SLT) can possibly be understood if there exists a new kind of probability measure, p_r, which drives thermodynamic processes and this thermodynamic probability evolves in such a manner that, whenever this thermodynamic probability changes its value when a system moves from one state to the next, it always takes higher values than the value it previously held – i.e. $dp_r \geq 0$, at all physical and material times. In a nutshell, thermodynamic events will at the very least, progressively evolve from a probabilistically less likely state – to a probabilistically more likely state. Such an evolution sequence is what is naturally expected from probability calculus anchored on common binary logic where natural systems are expected to steadily progress into their most likely state.

In the construction of our new ideas, naturally, we expected that this thermodynamic probability p_r, would turn out to be the usual Boltzmann probability, i.e.

$$p_r = Z^{-1} \exp\left(-E_r/k_{\mathrm{B}}T\right) ,$$

where p_r is the probability that for a system at temperature T, the microstate with energy E_r, will be occupied and Z is the partition function. As will be demonstrated in the sequel paper [5, hereafter Paper IV], this probability p_r, cannot be the usual Boltzmann probability, but a new kind of probability associated not with the occupation of the given microstate, but its evolution; where by evolution, it is understood to mean – moving or progression from its present state to a new state altogether.

Further on, in the paper [2, hereafter Paper II], we demonstrated that Liouville's theorem [6] can actually be viewed as a subtle statement of the *First Law of Thermodynamics* (FLT). This we did by defining the Liouville density function, $\delta\varrho$, in-terms of some new physical quantity, $\delta\mathcal{S}_{\mathrm{TD}}$, that we called the *thermodynamic phase* (or the *thermodynamic action*), i.e. $\delta\varrho = \exp(\delta\mathcal{S}_{\mathrm{TD}}/\hbar)$, where $\hbar$ is Planck's normalized constant. Furthermore, in Paper IV, we shall identify $\delta\varrho$ as the appropriate thermodynamic probability of evolution, that is, the thermodynamic probability responsible for the SLT.

In the present paper, we shall demonstrate that when cast as a probability measure, $\delta\varrho$ naturally yields the universally celebrated quantum mechanical uncertainty principle of Heisenberg [1]. In addition to Heisenberg's lower limiting (i.e. $\delta E\,\delta t \geq \hbar/2$ and $\delta p\,\delta x \geq \hbar/2$) uncertainty principle, we derive an upper limiting uncertainty principle – i.e. $\delta E\,\delta t \leq \hbar/2$ and $\delta p\,\delta x \leq \hbar/2$. As initially pointed out in [7], this upper limiting uncertainty principle strongly appears to describe non-ponderable tachyonic particles.

Without a doubt, Heisenberg's quantum mechanical uncertainty principle is certainly one of the most famous aspects of quantum mechanics and this very aspect of the theory is universally regarded as the most distinctive feature of the theory. It is a unique characteristic feature which makes quan-

tum mechanics differ radically from all classical theories of the physical world. For example, the uncertainty principle for position and momentum $\delta p \, \delta x \leq \hbar/2$ states that one cannot simultaneously assign exact values to the position and momentum of a physical system. Rather, these quantities can only be determined with some intrinsic, inherent and characteristic uncertainties that cannot – simultaneously – become arbitrarily small.

In its popular understanding, the Heisenberg uncertainty principle is assumed to be a principle to do with the accuracy in the results of measurements of physical variables such as momentum, position, energy, *etc*. Strictly speaking, this is not true. For example Millette [8] argues that the Heisenberg uncertainty principle arises from the dependency of momentum on wave number ($\boldsymbol{p} = \hbar\boldsymbol{k}$) that exists at the quantum level, and that ultimately the uncertainty principle is purely a relationship between the effective widths of Fourier transform pairs of conjugate variables. Our ideas propagated herein do support these views and as an addition, these quantum mechanical uncertainties associated with physical variables are seen to arise from pure stochastic processes occurring on some new phase-space that we have coined the stochastic phase space.

Now, in closing this introductory section, we shall give a synopsis of the present paper – i.e. this paper is organised as follows: in §2, we derive the uncertainty relations that govern ordinary ponderable matter and thereafter in §3, we derive the uncertainty relations that govern exotic non-ponderable matter. Lastly, in §4, we give a general discussion.

2 Derivation of the uncertainly principle

As stated in the introductory section, we are going to demonstrate in this section (i.e. in §2.2) that one can derive in a very natural and logical manner, the position-momentum and energy-time quantum mechanical Heisenberg uncertainty principle on the newly proposed *Stochastic Phase-Space* (hereafter $\delta\Gamma$-space) upon which we demonstrated [2] the naturally implied existence of the FLT from Liouville's theorem. In addition to Heisenberg's uncertainty principle, we will also derive in §3, upper limiting position-momentum and energy-time uncertainty principles and these upper limiting uncertainty principles describe non-ponderable tachyonic particles.

Before we proceed, we need to explain what it is we mean by *upper limiting uncertainty principle*. If there is an upper limiting uncertainty principle, from the viewpoint of common logic, there also must be a *lower limiting uncertainty principle*. Indeed, the uncertainty principle of Heisenberg is a lower limiting uncertainty principle because it gives the lowest possible value that the product of the energy (δE) & time (δt), and momentum (δp) & position (δr) uncertainties would ever take. That is to say, the products $\delta E \, \delta t$ and $\delta p \, \delta r$, can take whatever value they can or may take for so long as this value does not exceed the minimum threshold value of $\hbar/2$, hence, in this way, it becomes pristine clear that the Heisenberg un-

certainty principle ($\delta E \, \delta t \geq \hbar/2$ and $\delta p \, \delta r \geq \hbar/2$) is indeed a lower limiting uncertainty principle.

Now, if – by the sleight of hand, we are to flip the sign in the Heisenberg lower limiting uncertainty principle so that we now have $\delta E \, \delta t \leq \hbar/2$ and $\delta p \, \delta r \leq \hbar/2$, the resulting uncertainty principle is an upper limiting uncertainty principle since it now gives an upper limit in the value that the products ($\delta E \, \delta t$ and $\delta p \, \delta r$) of the uncertainties can ever take. Whence, we must hasten at this point and say we already have discussed the implications of a upper limiting uncertainty principle in our earlier works (i.e. in [7]) where we argued that if such particle exist to being with, not only will they travel at superluminal speeds – they also will have to be non-ponderable as-well; that is to say, they must be invisible and absolutely permeable. In simpler colloquial terms, such particles must be capable of passing through solid walls with no hindrance at all whatsoever.

2.1 Preliminaries

Now, before we can go on to present our derivation of Heisenberg's uncertainty principle in §2.2, we will need to set-up the stage for that event. First, in order for that, we shall give in §2.1.1, a description of the particle system that we shall consider, and, in §2.1.2, we shall describe the normalization across all spacetime for the thermodynamic probability function $\delta\varrho$ and in §2.1.3, we shall describe the normalization across a given space-and-momentum axis for the thermodynamic probability function, $\delta\varrho$. Lastly, in §2.1.4, we present some useful mathematical equations that we will need in our endeavours to derive the Heisenberg uncertainty principle.

2.1.1 Description of particle system

As initially suggested in Paper II, we envisage the existence of two mutually exclusive spacetimes and these we have termed – the *Classical Canonical Spacetime* (hereafter, CC-Spacetime), and, the *non-Canonical Spacetime* (hereafter, NC-Spacetime). The NC-Spacetime can also be called the *Stochastic Spacetime*. On the deterministic CC-Spacetime, a particle has its usual deterministic classical four position ($x, y, z, c_0 t$) that we are used to know, while on the non-deterministic NC-Spacetime, the non-deterministic jittery quantum randomness and fuzziness associated with the usual deterministic classical canonical position ($\delta x, \delta y, \delta z, c_0 \delta t$) are defined on this non-deterministic NC-Spacetime.

For example, considering only the x-axis, a particle will have x as its canonical position and δx as its associated non-canonical position as defined on the NC-Spacetime. It is δx that should give this particle the quantum fuzziness leading to the weird quantum probabilistic nature of physical systems. For the human observer – assuming zero human-induced error in measuring the position of the particle – the effective position $\hat{x}$ of the particle at any given time is $\hat{x} = x \pm \delta x$. So, in general, x^μ is the canonical four position of the particle and

δx^μ is the associated quantum randomness that leads to the mysterious, strange and bizarre fuzzy quantum probabilistic nature of natural systems.

In our description above, when we say particle, we mean a point-particle – i.e. particles of zero spatial dimension, hence zero volume. Obviously, there will be some trouble in accepting this – as point-particles, are – in physics – no more than idealization of real finite-sized particles that are smeared-out in a finite region of space. That is, a point-particle is in general an appropriate or convenient representation of any object whatever its size, shape, and structure – all these details of size, shape, and structure, *etc*, are irrelevant under the general particle model.

To further complicate this issue of the particle description of matter, we all know pretty well that the existence of point-particles is strictly forbade by Heisenberg's uncertainty principle. With this in mind, of these particles, what we envisage is them having all their charge such as their gravitational mass and electrical charge being concentrated on that very single point with this point being trapped in the finite sized spherical region of radius: $\delta r = \sqrt{(\delta x)^2 + (\delta y)^2 + (\delta z)^2}$, with the the centre of this finite spherical region fixed in space about the canonical position (x, y, z). Because of the fields that the trapped charged point-particle carries – i.e. fields with which this particle interacts with other particles; the fuzzy, random wandering and dotting back-and-forth, up-and-about of this particle inside this finite region should create the impression of a solid billiard-like ball of radius δr with oft cause the bulk of its charge (gravitational, electrical, *etc*) expected to be trapped in this spherical region. Surely, such a particle-system will be localized and it will have the property of ponderability that we experience with electrons, protons, *etc*. Let us call such a particle-system, a *Ponderable Material Particle*.

Now, for a minute, let us assume that the above described point-particle is not trapped. If that were the case, then, what is it that we are going to have for such a particle-system? Clearly, it must be an unbounded point-particle that is free to roam all of the Universe's length, breath and depth – from one end of the Universe, to the other in an instant! Such a particle-system should have its charge (gravitational, electrical, *etc*) spread-out evenly throughout the entire Universe. Not only this, while such a particle-system will have a definite fixed canonical position, the entire particle-system must be invisible as it will not have the property of ponderability (localization). Likewise, let us call such a particle-system, a *non-Ponderable Material Particle*.

Now, as shall soon become clear in our derivation of Heisenberg's uncertainty principle, two classes of particles will emerge and the first is that class whose random quantum fuzziness as described on the NC-Spacetime obeys the usual quantum mechanical uncertainty principle of Heisenberg, i.e. $\delta E \, \delta t \geq \hbar/2$ and $\delta p \, \delta r \geq \hbar/2$; and these particles travel at speeds

less than, or equal to the speed of light in *vacuo*. The second class is that of particles whose quantum fuzziness as described on the NC-Spacetime obeys not the usual quantum mechanical uncertainty principle Heisenberg, but obey the converse of Heisenberg's uncertainty principle, namely $\delta E \, \delta t \leq \hbar$ and $\delta p \, \delta r \leq \hbar$ and these particles travel at speeds that are at the very least, greater than the speed of light in *vacuo*.

At this juncture, we feel very strongly that we have prepared our reader to meet the strange new proposal of invisible particles that travel at superluminal speeds, thus – assuming the reader somewhat accepts or at the very least, finds some modicum of sense in what we have had to say above – we shall quietly proceed to the main business of this paper – that of demonstrating the natural existence of Heisenberg's uncertainty principle on the proposed NC-Spacetime where the jittery, fuzzy quantum randomness has here been defined.

2.1.2 Normalization across all space

If $\delta\varrho$ is assumed to be some probability function, then it must be normalizable. Normalization is oft cause one of the most fundamental and most basic properties that a probability function must satisfy. As is the norm: normalization of this function, $\delta\varrho$, across all of the six dimensions of $\delta\Gamma$-space requires that:

$$\frac{1}{\hbar^3} \overbrace{\int\int\int}^{\delta p_{\max}} \underbrace{\overbrace{\int\int\int}^{\delta r_{\max}}}_{\delta r_{\min}} (\delta\varrho_+) \, d^3x \, d^3p = 1, \qquad (1)$$

where: $\delta\varrho_+ = \delta\varrho_x^+ \delta\varrho_y^+ \delta\varrho_z^+ \delta\varrho_0^+$. In writing $\delta\varrho$ in (1), we have appended a subscript $+$ and this is not a mistake, it is deliberate. This $+$ appendage has been instituted – for latter purposes – so that a distinction can be made between a thermodynamic system with a positive $\delta\mathcal{S}_{TD}$ thermodynamic phase (action) and that with a negative $-\delta\mathcal{S}_{TD}$ thermodynamic phase (action), i.e.: $\delta\varrho_+ = \delta\varrho_+(\delta\mathcal{S}_{TD})$, while: $\delta\varrho_- = \delta\varrho_-(-\delta\mathcal{S}_{TD})$. The two functions describe two different kinds of phenomenon, namely $\delta\varrho_+$ describes ponderable matter as we know it, while $\delta\varrho_-$ describes some (exotic) non-ponderable (invisible) tachyonic matter. This shall be made clear as we go, hence the need to make a distinction of $\delta\varrho_+$ and $\delta\varrho_-$.

Now, the normalization in (1) is the probability of finding the particle in the spatial ($\hat{r}$) and momentum ($\hat{p}$) region:

$$\begin{aligned} \delta r_{\min} \leq \quad &\hat{r} \quad \leq \delta r_{\max} \\ \delta p_{\min} \leq \quad &\hat{p} \quad \leq \delta p_{\max} \, , \end{aligned} \qquad (2)$$

where $\hat{r}$ and $\hat{p}$ are the actual measured radial coordinate and magnitude of the momentum of the particle as measured from the spatial canonical point of origin of the particle (system) in question.

2.1.3 Normalization across a given axis

Now, given that $\delta\varrho_+ = \exp(\delta S_{\mathrm{TD}}/\hbar)$, where:

$$\delta S_{\mathrm{TD}} = \delta\boldsymbol{p}\cdot\delta\boldsymbol{r} - \delta E\delta t = \delta p_\mu \delta x^\mu, \qquad (3)$$

it follows that the quantities $\delta\varrho_x^+$, $\delta\varrho_y^+$, $\delta\varrho_z^+$, $\delta\varrho_0^+$ are such that:

$$
\begin{aligned}
\delta\varrho_x^+ &= \exp\left(\frac{\delta S_x}{\hbar}\right) &= \exp\left(\frac{\delta p_x \delta x}{\hbar}\right) &\quad\ldots\quad (a)\\
\delta\varrho_y^+ &= \exp\left(\frac{\delta S_y}{\hbar}\right) &= \exp\left(\frac{\delta p_y \delta y}{\hbar}\right) &\quad\ldots\quad (b)\\
\delta\varrho_z^+ &= \exp\left(\frac{\delta S_z}{\hbar}\right) &= \exp\left(\frac{\delta p_z \delta z}{\hbar}\right) &\quad\ldots\quad (c)\\
\delta\varrho_0^+ &= \exp\left(-\frac{\delta S_0}{\hbar}\right) &= \exp\left(-\frac{\delta E\delta t}{\hbar}\right) &\quad\ldots\quad (d)
\end{aligned}
\qquad (4)
$$

where oft cause $\delta S_x = \delta p_x \delta x$, $\delta S_y = \delta p_y \delta y$, $\delta S_z = \delta p_z \delta z$, and, $\delta S_0 = \delta p_0 \delta x_0 = \delta E\delta t$. Clearly, written in this manner, these functions $\delta\varrho_x^+$, $\delta\varrho_y^+$, $\delta\varrho_z^+$, $\delta\varrho_0^+$ are the thermodynamic probability evolution functions describing the particle across the δx-δp_x axis, δy-δp_y axis, δz-δp_z axis, and the δt-δE axis respectively.

The probability of finding the particle along the x-p_x, y-p_y, z-p_z and t-E axis respectively, in the region of its bounds is unity and this is expressed:

$$\frac{1}{\hbar}\int_{\delta p_{\min}}^{\delta p_{\max}}\int_{\delta r_{\min}}^{\delta r_{\max}}\left(\delta\varrho_x^+\right)dxdp_x = 1, \qquad (5)$$

$$\frac{1}{\hbar}\int_{\delta p_{\min}}^{\delta p_{\max}}\int_{\delta r_{\min}}^{\delta r_{\max}}\left(\delta\varrho_y^+\right)dydp_y = 1, \qquad (6)$$

$$\frac{1}{\hbar}\int_{\delta p_{\min}}^{\delta p_{\max}}\int_{\delta r_{\min}}^{\delta r_{\max}}\left(\delta\varrho_z^+\right)dzdp_z = 1, \qquad (7)$$

$$\frac{1}{\hbar}\int_{\delta E_{\min}}^{\delta E_{\max}}\int_{\delta t_{\min}}^{\delta t_{\max}}\left(\delta\varrho_0^+\right)dtdE = 1. \qquad (8)$$

Before we can deduce the Heisenberg uncertainty principle from the above equations (5)-(8), we shall lay down some necessary mathematical formulae.

2.1.4 Necessary mathematical equations

In our derivation of Heisenberg's uncertainty principle in §2.2 and §2.3, we are going to encounter the function e^{ax}/x, where x is the variable and a is some constant. Of this function, we will need to know its integral and limit as $x \mapsto 0$. It is not difficult to show that:

$$\int\left(\frac{e^{ax}}{x}\right)dx = \frac{e^{ax}}{ax} + k, \qquad (9)$$

where k is some integration constant and:

$$\lim_{x\to 0}\left(\frac{e^{ax}}{x}\right) = a. \qquad (10)$$

Now, we are ready to derive Heisenberg's uncertainty principle (and more).

2.2 Position-momentum uncertainty

In this section, we are now going to derive a lower and upper bound uncertainty principle for momentum and position. Taking (5) and substituting $\delta\varrho_x^+ = \exp(\delta p_x\delta x/\hbar)$, we will have:

$$\frac{1}{\hbar}\int_{\delta p_{\min}}^{\delta p_{\max}}\int_{\delta r_{\min}}^{\delta r_{\max}}\exp\left(\frac{\delta p_x\delta x}{\hbar}\right)dxdp_x = 1. \qquad (11)$$

Now, using the result of (9) to integrate (11) with respect to x, and evaluating the resulting integral, we will have:

$$\int_{\delta p_{\min}}^{\delta p_{\max}}\left(\frac{e^{\delta p_x\delta r_{\max}/\hbar} - e^{\delta p_x\delta r_{\min}/\hbar}}{\delta p_x}\right)dp_x = 1. \qquad (12)$$

Further, we need to integrate (12) with respect to p_x. In doing so, we will encounter again an integral of the form given in (9). The result of this integration is therefore:

$$\hbar\left[\frac{e^{\delta p_x\delta r_{\max}/\hbar}}{\delta p_x\delta r_{\max}} - \frac{e^{\delta p_x\delta r_{\min}/\hbar}}{\delta p_x\delta r_{\min}}\right]_{\delta p_{\min}}^{\delta p_{\max}} = 1. \qquad (13)$$

Evaluating this, we will have:

$$
\overbrace{\frac{\hbar e^{\delta p_{\max}\delta r_{\max}/\hbar}}{\delta p_{\max}\delta r_{\max}}}^{\textbf{Term I}} - \overbrace{\frac{\hbar e^{\delta p_{\min}\delta r_{\max}/\hbar}}{\delta p_{\min}\delta r_{\max}}}^{\textbf{Term II}}
$$

$$
-\underbrace{\frac{\hbar e^{\delta p_{\max}\delta r_{\min}/\hbar}}{\delta p_{\max}\delta r_{\min}}}_{\textbf{Term III}} + \underbrace{\frac{\hbar e^{\delta p_{\min}\delta r_{\min}/\hbar}}{\delta p_{\min}\delta r_{\min}}}_{\textbf{Term IV}} = 1. \qquad (14)
$$

Furthermore, for ponderable material particles, as discussed in §2.1.1, we want our particle system to be bounded (trapped) between the regions $0 \le \hat{r} \le \delta r_{\max}$ and $0 \le \hat{p}_x \le \delta p_{\max}$. This means that we must evaluate (14) in the limits $\delta r_{\min} \mapsto 0$ and $\delta p_{\min} \mapsto 0$.

Now, making use of the limit given (10), it follows that as:

$$
\begin{aligned}
\delta r_{\min} &\mapsto 0,\\
\delta p_{\min} &\mapsto 0,
\end{aligned}
\qquad (15)
$$

for the Terms I, II, III and IV in (14), we will have:

$$
\begin{aligned}
\text{Term I} &= \frac{\hbar e^{\dfrac{\delta p_{\max}\delta r_{\max}}{\hbar}}}{\delta p_{\max}\delta r_{\max}},\\
\text{Term II} &\mapsto 1,\\
\text{Term III} &\mapsto 1,\\
\text{Term IV} &\mapsto 1,
\end{aligned}
\qquad (16)
$$

hence from (16), it follows from this that (14) will reduce to:

$$\hbar e^{\delta p_{\max}\delta r_{\max}/\hbar}/\delta p_{\max}\delta r_{\max} - 1 = 1,$$

where, after some re-arrangement, we will have:

$$\frac{\frac{1}{2}\hbar}{\delta p_{max}\delta r_{max}} = e^{-\delta p_{max}\delta r_{max}/\hbar} \,. \tag{17}$$

From a meticulous inspection of (17), it is clear and goes without saying that in order for this equation to hold true $\delta p_{max}\delta r_{max} > 0$, hence:

$$\frac{\frac{1}{2}\hbar}{\delta p_{max}\delta r_{max}} = e^{-\delta p_{max}\delta r_{max}/\hbar} < 1 \,, \tag{18}$$

thus, we will have:

$$\delta p_{max}\delta r_{max} > \frac{1}{2}\hbar \,. \tag{19}$$

With the subscript "max" removed from p_{max} and r_{max}, this (19) is without any doubt whatsoever the famous 1927 position-momentum quantum mechanical uncertainty principle of Heisenberg. One can work this out for the other three cases – i.e. for the $(\delta y, \delta p_y)$ dimension as given in (6) and the $(\delta z, \delta p_z)$ dimension as given in (7) and they would arrive at the same result.

It is important to note that the exact Heisenberg upper uncertainty principle involves a *greater than or equal to sign*, that is "$\geq$", yet in (19), the equal sign "=" is missing. This issue shall be addressed in Paper IV where it shall be seen that this case represents only those particles that travel at the speed of light. Next, we consider the energy-time uncertainty relation.

2.3 Time-energy

Now, in §2.3.1 and §2.3.2, we are going to derive a lower and an upper bound uncertainty principle for energy and time and as we do this, we must have at the back of our mind that stable ponderable particles ought to have no upper bound in their temporal fluctuations. Yes, they can only have a lower bound in their temporal fluctuations and this lower bound must co-incide with the moment of their creation. On the contrary, unstable ponderable particles ought to have a finite upper bound in their temporal fluctuation.

2.3.1 Lower bound energy-time uncertainty

We are now going to derive the energy-time uncertainty principle. The derivation is similar to the one given in §2.2 above. To that end, from (4d) and (8), we know that:

$$\frac{1}{\hbar}\int_{\delta E_{min}}^{\delta E_{max}}\int_{\delta t_{min}}^{\delta t_{max}} \exp\left(-\frac{\delta E\delta t}{\hbar}\right)dtdE = 1 \,. \tag{20}$$

Now, using (9) to evaluate (20), we obtain the following:

$$\overbrace{\frac{\hbar e^{-\delta E_{max}\delta t_{max}/\hbar}}{\delta E_{max}\delta t_{max}}}^{\text{Term I}} - \overbrace{\frac{\hbar e^{-\delta E_{min}\delta t_{max}/\hbar}}{\delta E_{min}\delta t_{max}}}^{\text{Term II}}$$

$$-\underbrace{\frac{\hbar e^{-\delta E_{max}\delta t_{min}/\hbar}}{\delta E_{max}\delta t_{min}}}_{\text{Term III}} + \underbrace{\frac{\hbar e^{-\delta E_{min}\delta t_{min}/\hbar}}{\delta E_{min}\delta t_{min}}}_{\text{Term IV}} = 1 \,. \tag{21}$$

In the limit as:

$$\begin{aligned}\delta t_{min} &\mapsto 0 \,,\\ \delta E_{min} &\mapsto 0 \,,\end{aligned} \tag{22}$$

for Terms I, II, III and IV in (21), according to (10), we will have:

$$\begin{aligned}\text{Term I} &= \frac{\hbar e^{\frac{\delta E_{max}\delta t_{max}}{\hbar}}}{\delta E_{max}\delta t_{max}} \,,\\[6pt] \text{Term II} &\mapsto 1 \,,\\[6pt] \text{Term III} &\mapsto 1 \,,\\[6pt] \text{Term IV} &\mapsto 1 \,,\end{aligned} \tag{23}$$

hence, it follows from this – that (21) will reduce to:

$$\hbar e^{\delta E_{max}\delta t_{max}/\hbar}/\delta E_{max}\delta t_{max} - 1 = 1 \,,$$

where, after some algebraic re-arrangement, we can rewrite this equation as:

$$\frac{\frac{1}{2}\hbar}{\delta E_{max}\delta t_{max}} = e^{-\delta E_{max}\delta t_{max}/\hbar} \,. \tag{24}$$

Similarly, from an inspection of (24), one will clearly obtain that for this equation holds true $\delta E_{max}\delta t_{max} > 0$, hence:

$$\frac{\frac{1}{2}\hbar}{\delta E_{max}\delta t_{max}} = e^{-\delta E_{max}\delta t_{max}/\hbar} < 1 \,, \tag{25}$$

thus:

$$\delta E_{max}\delta t_{max} > \frac{1}{2}\hbar \,. \tag{26}$$

Once again, this is the famous 1927 energy-time quantum mechanical uncertainty principle of Heisenberg. Just as in (19), the reason for having the *greater than sign* and not the *greater than or equal to sign* are the same as those given in the case of (19). This uncertainty relation (i.e. (26)) describes a ponderable (spatially bound) material particle that is unstable and has a lifetime τ that is such that $\tau < \delta t_{max}$.

2.3.2 Upper bound energy-time uncertainty

Now, for the same reason given in §2.3.1, we are going to proceed further and consider the case of a ponderable material particle system that has no upper bound in its temporal fluctuations – i.e. a stable ponderable material particle system that can live forever (e.g. like an electron or a proton). Such a particle will have $\delta t_{\max}$ and $\delta E_{\max}$ being such that:

$$
\begin{aligned}
\delta t_{\max} &\mapsto \infty, \\
\delta E_{\max} &\mapsto \infty.
\end{aligned}
\tag{27}
$$

According to (10) under the given conditions (i.e. (27)), for the Terms I, II, III and IV in (21), we will have:

$$
\begin{aligned}
\text{Term I} &\mapsto 0, \\
\text{Term II} &\mapsto 0, \\
\text{Term III} &\mapsto 0, \\
\text{Term IV} &= \frac{\hbar e^{-\frac{\delta E_{\min}\delta t_{\min}}{\hbar}}}{\delta E_{\min}\delta t_{\min}},
\end{aligned}
\tag{28}
$$

hence, it follows from this that (21) will reduce to:

$$
\hbar e^{-\delta E_{\min}\delta t_{\min}/\hbar} / \delta E_{\min}\delta t_{\min} = 1,
$$

where, after some basic algebraic re-arrangement, we can rewrite this equation as:

$$
\frac{\hbar}{\delta E_{\min}\delta t_{\min}} = e^{\delta E_{\min}\delta t_{\min}/\hbar}.
\tag{29}
$$

As before, it is not difficult to see that for (29) to hold true, this requires that $\delta E_{\min}\delta t_{\min} > 0$, hence, and from this, it clearly follows that:

$$
\frac{\hbar}{\delta E_{\min}\delta t_{\min}} = e^{\delta E_{\min}\delta t_{\min}/\hbar} < 1,
\tag{30}
$$

thus:

$$
\delta E_{\min}\delta t_{\min} < \hbar.
\tag{31}
$$

Insofar as its interpretation is concerned, by no stitch of the imagination is this (31) related to the famous 1927 energy-time quantum mechanical uncertainty principle of Heisenberg and this is so because of the *less-than-sign* "<" appearing in it. What this equation is "telling" us is that the energy and time fluctuations are not bound above, but below. When it comes to the lifetime of the particle in question, this translates to the reality that the particle can live forever – i.e. $\tau = \infty$. Therefore, this uncertainty relation describes stable ponderable particle systems – i.e. ordinary electrons and protons, which by-and-large strongly appear to be stable particle systems.

3 Non-ponderable matter

From a symmetry and *bona fide* mathematical standpoint, if we have the physics of particles described by the thermodynamic phase $+\delta S_{\mathrm{TD}}$, there surely is nothing wrong, but everything natural and logical for one to consider the physics of particle systems described by the opposite thermodynamic phase – i.e. $-\delta S_{\mathrm{TD}}$. Such necessary and beautiful symmetry considerations is what lead the great English theoretical physicist – Paul Adrian Maurice Dirac (1902-1984) to foretell the existence of antimatter [9–11]. We here consider the said particle systems whose thermodynamic phase is $-\delta S_{\mathrm{TD}}$.

Before even going into investigating the said particle systems, natural questions will begin to flood the mind, questions such as: *Will such particles violate the FLT?* The answer is: *No, they will not.* To see this, one simply substitutes $-\delta S_{\mathrm{TD}}$ into the equations of Paper II, where-from they certainly will come to the inescapable conclusion that these particles will indeed obey the FLT. Further – a question such as: *Will these particle systems violate the SLT?* may also visit the curious and searching mind. An answer to this will be provided in Paper IV.

Furthermore – in the extreme and zenith of one's state of wonderment, they might excogitate: *Will such particles be visible and ponderable?* By visible it is understood to mean: will these particle systems emit or reflect electromagnetic radiation that we are able to sense? And by ponderable, we mean will such particle systems be able to clump-up and form touchable materials like rocks, *etc*? This is the question we are going to answer. To preempt our findings, such particle systems will be invisible and non-ponderable.

To commence our expedition, we shall start by writing down the functions $\delta\varrho_x^-, \delta\varrho_y^-, \delta\varrho_z^-, \delta\varrho_0^-$ and these are such that:

$$
\begin{aligned}
\delta\varrho_x^- &= \exp\left(-\frac{\delta S_x}{\hbar}\right) &= \exp\left(-\frac{\delta p_x \delta x}{\hbar}\right) &\quad \dots \quad (a) \\
\delta\varrho_y^- &= \exp\left(-\frac{\delta S_y}{\hbar}\right) &= \exp\left(-\frac{\delta p_y \delta y}{\hbar}\right) &\quad \dots \quad (b) \\
\delta\varrho_z^- &= \exp\left(-\frac{\delta S_z}{\hbar}\right) &= \exp\left(-\frac{\delta p_z \delta z}{\hbar}\right) &\quad \dots \quad (c) \\
\delta\varrho_0^- &= \exp\left(\frac{\delta S_0}{\hbar}\right) &= \exp\left(\frac{\delta E \delta t}{\hbar}\right) &\quad \dots \quad (d)
\end{aligned}
\tag{32}
$$

Now, just as in the case of ponderable matter in the previous section, in order for us to derive the implied uncertainty relations from (32), we are going to consider (in §3.1, §3.2 and §3.3, respectively) the normalization of $\delta\varrho_x^-$ and $\delta\varrho_0^-$.

3.1 Lower bound position-momentum uncertainty

As before, normalization of $\delta\varrho_x^-$ requires that:

$$
\frac{1}{\hbar}\int_{\delta p_{\min}}^{\delta p_{\max}}\int_{\delta r_{\min}}^{\delta r_{\max}}\exp\left(-\frac{\delta p_x \delta x}{\hbar}\right)dx\,dp_x = 1.
\tag{33}
$$

Just as we have already done with (11) and (20); integrating and evaluating (33), we obtain:

$$
\overbrace{\frac{\hbar e^{-\delta p_{max}\delta r_{max}/\hbar}}{\delta p_{max}\delta r_{max}}}^{\textbf{Term I}} - \overbrace{\frac{\hbar e^{-\delta p_{min}\delta r_{max}/\hbar}}{\delta p_{min}\delta r_{max}}}^{\textbf{Term II}}
$$

$$
- \underbrace{\frac{\hbar e^{-\delta p_{max}\delta r_{min}/\hbar}}{\delta p_{max}\delta r_{min}}}_{\textbf{Term III}} + \underbrace{\frac{\hbar e^{-\delta p_{min}\delta r_{min}/\hbar}}{\delta p_{min}\delta r_{min}}}_{\textbf{Term IV}} = 1. \quad (34)
$$

Likewise, with (34) in place, one may try to bound the particle in space and momentum, in much the same way as it has been done in §2.2 by instituting the asymptotic conditions $\delta r_{min} \mapsto 0$ and $\delta p_{min} \mapsto 0$. So doing, they surely would obtain the unpleasant result:

$$
\hbar e^{-\delta p_{max}\delta r_{max}/\hbar}/\delta p_{max}\delta r_{max} = 0.
$$

This result is surely unpleasant because it means that we must have $\delta p_{max}\delta r_{max} = \infty$. Overall, this means that this particle system has no upper bounds in quantum of action $\delta p_{max}\delta r_{max}$; this surely is uncomfortable as the quantum of action must be bound either above or below. Given this uncomfortable result $\delta p_{max}\delta r_{max} = \infty$, a much better way to approach this particle system is to start off by setting no upper bounds in space and momentum and in the end obtain finite lower bounds in the quantum of action $\delta E \delta t$, that is to say, start off by setting:

$$
\begin{aligned}
\delta r_{max} &\mapsto \infty, \\
\delta p_{max} &\mapsto \infty.
\end{aligned} \quad (35)
$$

Instituting the above (35) limits into (34), for the Terms: (I), (II), (III) and (IV), one obtains:

$$
\begin{aligned}
\text{Term I} &\mapsto 0, \\
\text{Term II} &\mapsto 0, \\
\text{Term III} &\mapsto 0, \\
\text{Term IV} &= \frac{\hbar e^{-\frac{\delta p_{min}\delta r_{min}}{\hbar}}}{\delta p_{min}\delta r_{min}},
\end{aligned} \quad (36)
$$

hence:

$$
\frac{\hbar e^{-\frac{\delta p_{min}\delta r_{min}}{\hbar}}}{\delta p_{min}\delta r_{min}} = 1. \quad (37)
$$

In much the same fashion as in the preceding sections, re-arranging this (37), we will have:

$$
\frac{\hbar}{\delta p_{min}\delta r_{min}} = e^{\frac{\delta p_{min}\delta r_{min}}{\hbar}} > 1, \quad (38)
$$

hence:

$$
\delta p_{min}\delta r_{min} < \hbar. \quad (39)
$$

This means the fuzziness in the momentum and spatial location of the particle about its canonical centre is bounded above and not below.

3.2 Lower bound energy-time uncertainty

Further, for the energy-time uncertainty relation, normalization of $\delta \varrho_0^-$ requires that:

$$
\frac{1}{\hbar} \int_{\delta E_{min}}^{\delta E_{max}} \int_{\delta t_{min}}^{\delta t_{max}} \exp\left(\frac{\delta E \delta t}{\hbar}\right) dt\, dE = 1. \quad (40)
$$

As before, integrating and evaluating this (40), we obtain:

$$
\overbrace{\frac{\hbar e^{-\delta E_{max}\delta t_{max}/\hbar}}{\delta E_{max}\delta t_{max}}}^{\textbf{Term I}} - \overbrace{\frac{\hbar e^{-\delta E_{min}\delta t_{max}/\hbar}}{\delta E_{min}\delta t_{max}}}^{\textbf{Term II}}
$$

$$
- \underbrace{\frac{\hbar e^{-\delta E_{max}\delta t_{min}/\hbar}}{\delta E_{max}\delta t_{min}}}_{\textbf{Term III}} + \underbrace{\frac{\hbar e^{-\delta E_{min}\delta t_{min}/\hbar}}{\delta E_{min}\delta t_{min}}}_{\textbf{Term IV}} = 1. \quad (41)
$$

In the limit as:

$$
\begin{aligned}
\delta t_{max} &\mapsto \infty, \\
\delta E_{max} &\mapsto \infty,
\end{aligned} \quad (42)
$$

for the Terms I, II, III and IV in (41), we will have:

$$
\begin{aligned}
\text{Term I} &\mapsto 0, \\
\text{Term II} &\mapsto 0, \\
\text{Term III} &\mapsto 0, \\
\text{Term IV} &= \frac{\hbar e^{-\frac{\delta E_{min}\delta t_{min}}{\hbar}}}{\delta E_{min}\delta t_{min}},
\end{aligned} \quad (43)
$$

hence, it follows from this that (41) will reduce to:

$$
\hbar e^{-\delta E_{min}\delta t_{min}/\hbar}/\delta E_{min}\delta t_{min} = 1,
$$

where, after some re-arrangement, we can rewrite:

$$
\frac{\hbar}{\delta E_{min}\delta t_{min}} = e^{\delta E_{min}\delta t_{min}/\hbar}. \quad (44)
$$

As before, from a meticulous inspection of (44), it is abundantly clear that $\delta E_{min}\delta t_{min} > 0$, hence:

$$
\frac{\hbar}{\delta E_{min}\delta t_{min}} = e^{\delta E_{min}\delta t_{min}/\hbar} < 1, \quad (45)
$$

thus:

$$
\delta E_{min}\delta t_{min} < \hbar. \quad (46)
$$

Just as with (39), (46) means that the fuzziness in the energy and temporal fluctuations of the particle are bounded above and not below.

3.3 Upper bound energy-time uncertainty

Lastly, we now consider the case of a non-ponderable material particle system that has no upper bound in its temporal fluctuation – i.e. a stable non-ponderable material particle system that can live forever. Such a particle will have $\delta t_{\max}$ and $\delta E_{\max}$ such that:

$$
\begin{aligned}
\delta t_{\max} &\mapsto \infty, \\
\delta E_{\max} &\mapsto \infty.
\end{aligned}
\tag{47}
$$

Under the given conditions (i.e. (47)), for the Terms I, II, III and IV in (41), we will have:

$$
\begin{aligned}
\text{Term I} &\mapsto 0, \\
\text{Term II} &\mapsto 0, \\
\text{Term III} &\mapsto 0, \\
\text{Term IV} &= \frac{\hbar e^{-\frac{\delta E_{\min}\delta t_{\min}}{\hbar}}}{\delta E_{\min}\delta t_{\min}},
\end{aligned}
\tag{48}
$$

hence, it follows from this that (21), will reduce to:

$$
\hbar e^{-\delta E_{\min}\delta t_{\min}/\hbar}/\delta E_{\min}\delta t_{\min} = 1,
$$

where, after some re-arrangement, we can rewrite:

$$
\frac{\hbar}{\delta E_{\min}\delta t_{\min}} = e^{\delta E_{\min}\delta t_{\min}/\hbar}.
\tag{49}
$$

Likewise, for it to hold true always, (49) requires that $\delta E_{\min}\delta t_{\min} < 0$, hence:

$$
\frac{\hbar}{\delta E_{\min}\delta t_{\min}} = e^{\delta E_{\min}\delta t_{\min}/\hbar} < 1,
\tag{50}
$$

thus:

$$
\delta E_{\min}\delta t_{\min} < \hbar.
\tag{51}
$$

Again, we here have an upper bounded uncertainty relation.

4 General discussion

Since the inception of Heisenberg's uncertainty principle in 1927, several attempts see e.g. [8,12–15, and references therein] have been made to derive this mysterious mathematical relationship from much more fundamental soils of physics than those on which Heisenberg [1] derived this relation. In his original paper, Heisenberg began by deriving the uncertainty relation for position and momentum on the basis of a supposed experiment in which an electron is observed using a γ-ray microscope and second, by consideration of the theory of the Compton effect, he proceeded to argue that the precision of the determination of position and momentum are connected by the uncertainty relation.

In 1929, using the usual definition of expectation values (inner product) of Hermitian Hilbert-space operators (observables) and the mathematical property of the *Cauchy–Bunyakovsky–Schwarz inequality*, Robertson [12] proceeded in a rigorous manner, to demonstrate a more general and fundamental origin of the quantum mechanical uncertainty principle. The present attempt is just but one such derivation – albeit – on the soils of a new kind of phase space – the Stochastic Phase Space.

However, unlike all previous attempts on the derivation of the uncertainty principle, what makes the present attempt different is that we have not only derived the lower limit uncertainty principle, but an upper bound uncertainty principle that seems to describe invisible non-ponderable particles that travel at superlumical speeds. This unique prediction seem to suggest not only the existence of darkmatter, but darkenergy as well. Dark matter is already required by physicists in order to explain the flat rotation curves of spiral galaxies, while dark energy is required to explain the supposed accelerated expansion of the Universe. This subject of invisible non-ponderable particles, dark matter and dark energy would require a separate and lengthy paper in order to cover it in a just manner.

Another important point to note about the present derivation is that the enigmatic jittery quantum randomness leading to the uncertainty principle is here an intrinsic and inherent property of all quantum mechanical systems, it (i.e. the jittery quantum randomness) is not induced by the act of measurement as is the case of Heisenberg's uncertainty principle and its latter versions or attempts at a derivation of this relation. Yes, human measure will introduce statistical errors that are statistically predictable. The stochastic quantum randomness is not predictable at all – not even by the most rigours known (or unknown, or yet to be unknown) statistical methods.

In closing, allow us to say that we have always held central to our philosophy of Physics the strong and seemingly unshakeable belief system similar to that of Albert Einstein – namely, that the fundamental laws of Nature are exact, and as such, one day it will be shown that this is the case. That is to say, in the character of Einstein's philosophy, we have held fast to his influential and deep philosophy that indeed *God does not play dice with the World*, and that *The moon exists whether or not one is looking at it or not*. Contrary to this, we must admit and say that as we continue to peer deeper into the fabric and labyrinth of physical and natural reality as it lies bare for us to marvel at, this dream or belief system now stands shattered into minuscule pieces – for it now seems clearer to us that the enigmatic jittery quantum randomness must be real.

References

1. Heisenberg W. K. Ueber den anschaulichen Inhalt der Quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 1927,

v. 43 (3), 172–198. English Translation: Wheeler J. A. and Zurek W. H., eds. Quantum Theory and Measurement, Princeton University Press, Princeton, NJ, 1983, pp. 62–84.

2. Nyambuya G. G. Liouville's Theorem as a Subtle Statement of the First Law of Thermodynamics. *Progress in Physics*, 2019, v. 15 (3), 178–181.

3. Heisenberg W. K. Physics and Beyond. George Allen & Unwin, 1st Edition, 1971.

4. Nyambuya G. G. A Simple Proof of the Second Law of Thermodynamics. *Progress in Physics*, 2019, v. 15 (3), 171–177.

5. Nyambuya G. G. On the Uni-Directional Thermodynamic Arrow of Evolution of Thermodynamic Systems. In Preparation.

6. Liouville J. Sur la Théorie de la Variation des Constantes Arbitraires. *Journ. de Math.*, 1838, v. 3, 342–349.

7. Nyambuya G. G. On the Plausibility of an Upper Bound Uncertainty Principle. *Prespacetime Journal*, 2014, v. 5 (10), 1018–1026.

8. Millette P. A. The Heisenberg Uncertainty Principle and the Nyquist-Shannon Sampling Theorem. *Progress in Physics*, 2013, v. 9 (3), 9–14.

9. Dirac P. A. M. The Quantum Theory of the Electron. *Proc. Roy. Soc. (London)*, 1928a, v. A117, 610–612.

10. Dirac P. A. M. The Quantum Theory of the Electron II. *Proc. Roy. Soc. (London)*, 1928b, v. A118, 351–361.

11. Dirac P. A. M. A Theory of Electrons and Protons. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 1930, v. 126 (801), 360–365.

12. Robertson H. P. The Uncertainty Principle. *Physical Review*, 1929, v. 34 (1), 163–164.

13. Davidson E. R. On Derivations of the Uncertainty Principle. *The Journal of Chemical Physics*, 1965, v. 42 (4), 1461–1462.

14. Briggs J. S. A Derivation of the Time-Energy Uncertainty Relation. *Journal of Physics: Conference Series*, 2008, 012002, 99.

15. Briggs J. S. and Rost J. M. *Foundations of Physics*, 2001, v. 31 (4), 693–712.

Magnetic Energy, Superconductivity, and Dark Matter

Hanno Essén

Department of Engineering Mechanics, Royal Institute of Technology, Osquars Backe 18, 100 44, Stockholm, Sweden.
E-mail: hanno@mech.kth.se

Magnetism due to the translational, possibly oscillatory, motion of charge, as opposed to the ordering of dipoles, is not well understood, but is well described by the Darwin Lagrangian. The Coulomb interaction is used universally in atomic, molecular and solid state physics, but its natural extension when going to higher accuracy, the magnetic Darwin-Breit interaction, is not. This interaction is a velocity dependent long range interaction and as such unfamiliar to the majority of theoreticians. The $(v/c)^2$ dependence makes it at most a perturbation in few-body systems, but does not stop it from becoming potentially important as the number of particles increase. For systems where particle velocities are correlated (or coherent) over larger distances this interaction is shown to have major consequences. Based on these findings I suggest that this interaction should be investigated as the interaction responsible for superconductivity. I also speculate that, on an interstellar scale, it is responsible for the missing dark matter. Some numerical estimates and intuitive arguments are presented in support, but no proofs. Instead it is my hope that the ideas presented will deserve further serious study.

A man hears what he wants to hear and disregards the rest.

Paul Simon in The Boxer

1 Introduction

We first introduce the Darwin Lagrangian which describes the magnetic interaction energy between moving charged particles. This is a velocity dependent long range interaction which is very small for few-body systems but which can become dominating in macroscopic systems. In particular the Lagrangian predicts that the effective mass, or equivalently inductive inertia, can grow with the square of the number of particles.

The Darwin Lagrangian makes simple predictions for particles that are assumed to have the same velocity. Here we use this constraint to study the effect of the magnetic interaction energy for collectively moving charges. The crucial fact that emerges from these studies is that the effective mass of many collectively moving particles far exceeds the sum of their rest masses. In the case of superconductivity this means that the zero-point energy of coherent oscillators decreases with the number of oscillators, and this presumably leads to the superconducting phase transition. In the case of cosmic plasma filaments it leads to the conclusion that their gravitational mass can far exceed the rest mass content of the participating particles. Could this be the missing dark matter? Some numerical estimates indicate that this is a possibility.

2 The Darwin Lagrangian

The Darwin Lagrangian [1] describes the majority of electromagnetic phenomena correctly. The exception is radiation, which is neglected. The theory behind this Lagrangian is presented in a few textbooks such as Landau and Lifshitz [2, §65] and Jackson [3, Sec. 12.6]. More extensive discussions

can be found in Page and Adams [4, Sec. 96], Podolsky and Kunz [5, Sec. 27], Szasz [6, Appendix], Schwinger *et al.* [7, Eq. (33.23)], or Stefanovich [8]. Basic articles of interest are Breitenberger [9], Kennedy [10], Essén [11–13]. Various applications of the Darwin Lagrangian illustrating its usefulness can be found in Kaufman [14], Stettner [15], Boyer [16, 17], Krause *et al.* [18], Essén *et al.* [19–25].

Vector potentials are not always mentioned in connection with the Darwin Lagrangian, but it can be derived by approximating the Liénard-Wiechert potentials. Landau and Lifshitz [2, §65] make a gauge transformation to the Coulomb gauge after truncating series expansions of these. Jackson [3, Sec. 12.6] solves the vector Poisson equation obtained by neglecting the time derivative in the wave equation. Page and Adams derive it by approximating the forces [4, Sec. 96]. It can also be motivated as the best approximately relativistic action-at-a-distance Lagrangian [10, 26] and it can be shown to take retardation into account to order $(v/c)^2$.

The Darwin Lagrangian for N charged particles, of mass m_a and charge e_a, can be written

$$L_D = \sum_{a=1}^{N} \left[\frac{m_a}{2} \boldsymbol{v}_a^2 - \frac{e_a}{2} \phi_a(\boldsymbol{r}_a) + \frac{e_a}{2c} \boldsymbol{v}_a \cdot \boldsymbol{A}_a(\boldsymbol{r}_a) \right] \quad (1)$$

where

$$\phi_a(\boldsymbol{r}_a) = \sum_{b(\neq a)}^{N} \frac{e_b}{|\boldsymbol{r}_a - \boldsymbol{r}_b|} \quad (2)$$

and

$$\boldsymbol{A}_a(\boldsymbol{r}_a) = \sum_{b(\neq a)}^{N} \frac{e_b}{2c} \frac{[\boldsymbol{v}_b + (\boldsymbol{v}_b \cdot \hat{\boldsymbol{e}}_{ab})\hat{\boldsymbol{e}}_{ab}]}{|\boldsymbol{r}_a - \boldsymbol{r}_b|} . \quad (3)$$

Here $\hat{\boldsymbol{e}}_{ab} = (\boldsymbol{r}_a - \boldsymbol{r}_b)/|\boldsymbol{r}_a - \boldsymbol{r}_b|$, and relativistic corrections to the kinetic energy are neglected. In many circumstances

one can neglect the magnetic interaction energies since the Coulomb electric interaction dominates strongly, especially in few-body systems. As will be seen below, however, when there are macroscopic numbers of correlated charged particles this is no longer permissible. It is noteworthy that macroscopic numbers of correlated charged particles is the rule rather than an exception in plasmas, conductors, and superconductors.

3 Plasma oscillations

One can use (1) to calculate how a charge density of electrons oscillates relative to a fixed background of positive charge. For collective motion of N electrons with velocity $\boldsymbol{v} = \dot{x}\hat{\boldsymbol{e}}_x$ the kinetic energy is simply $T = Nm_e\dot{x}^2/2$. If one further assumes that the particles have fixed distributions in space apart form the relative translational motion one can get (nearly) analytical results for the remaining two terms, for simple geometries in the continuum limit. If we denote the displacement of the negative charges by x the total Coulomb potential energy is well approximated by,

$$\Phi(x) = \sum_{a=1}^{N_{tot}} \frac{e_a}{2} \phi_a(\boldsymbol{r}_a) = \Phi(0) + \frac{1}{2}\left(\frac{d^2\Phi}{dx^2}\right)_{x=0} x^2, \qquad (4)$$

in the limit of small x. Here $\Phi(0)$ is a large negative constant that does not contribute to the dynamics; the positive background only provides the restoring force in the oscillation. The assumption that the electrons ($m_a = m_e$, $e_a = -e$) move collectively along the x-direction simplifies the magnetic contribution, the third term in (1). One finds

$$
\begin{aligned}
U_D &= \sum_{a=1}^{N} \frac{e}{2c} \boldsymbol{v}_a \cdot \boldsymbol{A}_a(\boldsymbol{r}_a) \\
&= \left(\frac{e^2}{2c^2} \sum_{a=1}^{N-1}\sum_{b=a+1}^{N} \frac{1+\cos^2\theta_{ab}}{|\boldsymbol{r}_a - \boldsymbol{r}_b|}\right)\dot{x}^2,
\end{aligned}
\qquad (5)
$$

where $\cos\theta_{ab} = \hat{\boldsymbol{e}}_x \cdot \hat{\boldsymbol{e}}_{ab}$. For a charge density of electrons with fixed geometry this is simply a constant times $\dot{x}^2$. We thus find that the Darwin Lagrangian for the system becomes

$$L_D = N\left(\frac{1}{2}m_{\text{eff}}\dot{x}^2 - \frac{1}{2}\kappa x^2\right). \qquad (6)$$

Here m_{eff} is m_e plus a contribution from (5).

Calculations of the constants m_{eff} and κ can be done by elementary methods. The result will be a formula for the square of the oscillation frequency $\omega^2 = \kappa/m_{\text{eff}}$. This was done for a sphere of radius R in [19] with the result

$$\omega^2 = \frac{\frac{Ne^2}{R^3}}{m_e\left(1 + \frac{4}{5}\frac{r_e N}{R}\right)}. \qquad (7)$$

Here $r_e = e^2/(m_e c^2)$ is the classical electron radius. $\omega(R)$ is plotted in Fig. 1. In the limit of few particles, or negligible Nr_e/R, this gives the plasma oscillation frequency as
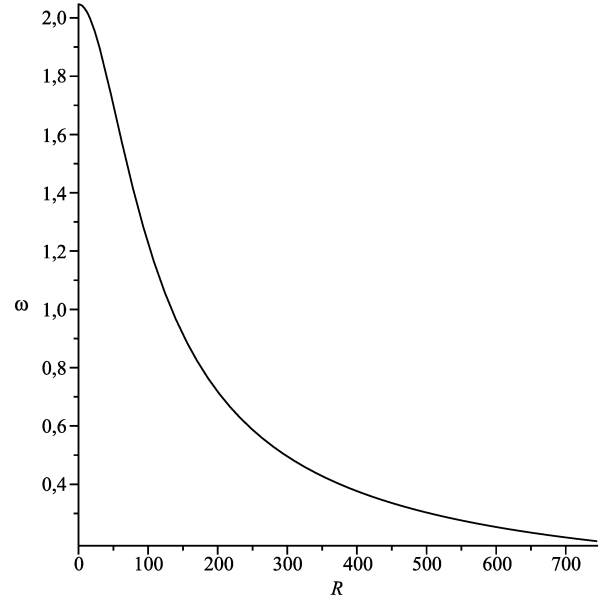
Fig. 1: The frequency ω of (7) as a function of radius R. Atomic units are used ($e = m_e = \hbar = 1$, $c = 137$) and the density is assumed to be one electron per sphere of one Bohr radius a_0. The formula plotted is $\omega(R) = \sqrt{4\pi/3\left(1 + \frac{16\pi}{15(137)^2}R^2\right)}$ and R is in atomic units (Bohr radii). The frequency is reduced by one order of magnitude at $R = 750\,a_0$.

normally given in the literature,

$$\omega_p^2 = \frac{4\pi}{3}\frac{e^2 n_0}{m_e} \qquad (8)$$

where $n_0 = N/V$ is the number density inside the sphere. In the opposite limit of macroscopic numbers of electrons N one obtains

$$\omega_\infty^2 = \frac{5c^2}{4R^2}. \qquad (9)$$

This seems to be the frequency of a longitudinal electromagnetic wave in the sphere. A similar calculation for a (two-dimensional) square of side length L gives a similar result,

$$\omega^2 = \frac{\frac{2e^2 N}{L^3}K_s}{m_e\left(1 + \frac{3}{4}\frac{Nr_e}{L}C_s\right)}, \qquad (10)$$

where $C_s = (4/3)\left[1 - \sqrt{2} - 3\ln\left(\sqrt{2}-1\right)\right]$ and $K_s = 16\,(2 - \sqrt{2})$ [27].

4 Superconductivity

In the early history of superconductivity it was conjectured that a transition of the electrons at the Fermi surface to a Wigner crystal [28] was responsible for the phase transition. Since no new interaction comes into play this did not seem correct, even if the Wigner crystal idea is still investigated

Hanno Essén. Magnetic Energy, Superconductivity, and Dark Matter

[29, 30]. When one takes the magnetic interaction energy into account, however, the zero point energy $E_0 = \hbar\omega/2$ and oscillation frequency of the (pairs of) electrons go down considerably if they oscillate coherently with coherence length R, as indicated in Fig. 1. It is interesting to note that Vasiliev [31, 32] finds that superconductivity is caused by ordering of the zero point oscillations. Frenkel [33] advanced the theory that the increased inductive inertia of correlated conduction electrons explains superconductivity, and the present author presented estimates indicating that the Darwin energy is important in superconductors [34]. In Fig. 1 it is seen that the zero point energy goes down by one order of magnitude in 750 Bohr-radii, assuming one electron per cubic Bohr-radius. In general coherence lengths in superconductors is one or two orders of magnitude larger [35], so the numbers are quite reasonable. The isotope effect agrees well with the assumption that lattice oscillations destroy the coherence.

5 Dark matter?

The decay time of currents is $\tau \sim \mathcal{L}/\mathcal{R}$ where $\mathcal{L}$ is inductance and $\mathcal{R}$ resistance. As emphasized by Kulsrud [36] these times are enormous in astrophysical plasmas. The currents producing astrophysical magnetic fields will only decay on a time scale comparable to the age of the universe. These plasmas are thus effectively superconducting. The effective mass m_{eff} of (6) is a measure of the inductance, or inductive inertia. Simple estimates show that this mass is in general much larger than the rest mass. That this is the case for conduction electrons in a metal was noted already in 1936 by Darwin [37] and several times later [23, 38].

It is tempting to speculate that dark matter is in fact due to magnetic energy in interstellar plasmas. Here we make some simple estimates. The Darwin magnetic energy, the first term of (6), $U_D = N m_{\text{eff}} \dot{x}^2/2$, will contribute $M_D = U_D/c^2$ to gravitational mass in the universe. Consider a cube of side length L. If we assume that the number of protons in this cube is N and that L also is a typical distance between them we find from (5) that

$$\frac{U_D}{c^2} = M_D \sim \frac{1}{4}\frac{e^2}{c^2}\frac{N^2}{L}\beta^2 \tag{11}$$

where $\beta = |\dot{x}|/c$. This magnetic mass should be compared to the total proton mass $M_p = N m_p$. The ratio is

$$\frac{M_D}{M_p} \sim \frac{(e^2/c^2)(N/L)}{4 m_p}\beta^2. \tag{12}$$

Putting in the numerical values gives

$$\frac{M_D}{M_p} \sim (3.83\times 10^{-19}\,\text{m})\frac{N}{L}\beta^2. \tag{13}$$

The number of protons is $N = n_p L^3$ where n_p is the proton number density. This gives

$$\frac{M_D}{M_p} \sim (3.83\times 10^{-19}\,\text{m})\,n_p\,L^2\beta^2. \tag{14}$$

To get some numbers we assume that $n_p = 4.0\,\text{m}^{-3}$ and that the ratio M_D/M_p is 10 (magnetic mass is 10 times proton mass). This gives

$$10 \sim 3.83\times 10^{-19}\,\times 4.0\,L^2\beta^2\,\text{m}^{-2}. \tag{15}$$

The side length of the cube over which velocity must be correlated is then

$$L \sim 2.5\times 10^9\,\beta^{-1}\,\text{m}. \tag{16}$$

assuming that the speed is $c/100$, so that $\beta = 10^{-2}$, we find that $L \sim 2.5\times 10^{11}$ m. This is somewhat more that one astronomical unit (AU $\approx 1.5\times 10^{11}$ m), a tiny distance in the interstellar perspective. So, with a density of 4 protons per cubic meter and a correlated speed of 1% of the speed of light over a distance of order of magnitude one AU one finds that the gravitational mass M_D of the magnetic energy is ten times the total proton rest mass. This suggests to me that dark matter may, in fact, reside in magnetic energy and the effective mass of the cosmic magnetic fields.

6 Conclusions

Since Darwin's 1936 paper [37] it should have been clear that investigations of conduction electrons in metals that do not take into account the magnetic interaction energy are meaningless. No amount of mathematical wizardry will make this interaction go away. It is also a natural candidate for emergent properties in larger systems, such as superconductivity, while remaining a perturbation in few body systems.

The insight that large plasmas with coherent velocities have energies that are many orders of magnitude larger than that corresponding to the rest mass of the constituent particles should be investigated as a possible candidate for dark matter. Recently Nicastro *et al.* [39] found that missing baryons are believed to reside in large-scale filaments in the warm-hot intergalactic medium. Perhaps the rest of the missing dark matter is also there in the form of magnetic energy?

References

1. Darwin C. G. The dynamical motions of charged particles. *Philos. Mag. (UK)*, 1920, v. 39, 537–551.

2. Landau L. D. and Lifshitz E. M. The Classical Theory of Fields, 4th edition. Pergamon, Oxford, 1975.

3. Jackson J. D. Classical Electrodynamics, 3rd edition. John Wiley, New York, 1999.

4. Page L. and Adams N. I. Electrodynamics. Van Nostrand, New York, 1940.

5. Podolsky B. and Kunz K. S. Fundamentals of Electrodynamics. Marcel Dekker, New York, 1969.

6. Szasz L. The Electronic Structure of Atoms. John Wiley, New York, 1992.

7. Schwinger J., DeRaad L. L. Jr., Milton K. A., and Tsai W.-Y. Classical Electrodynamics. Perseus Books, Reading, Massachusetts, 1998.

8. Stefanovich E. V. Relativistic Quantum Theory of Particles. II: A Non-Traditional Perspective on Space, Time, Particles, Fields, and Action-at-a-Distance. LAP Lambert, Saarbrücken, Germany, 2015.

9. Breitenberger E. Magnetic interactions between charged particles. *Am. J. Phys.*, 1968, v. 36, 505–515.

10. Kennedy F. J. Approximately relativistic interactions. *Am. J. Phys.*, 1972, v. 40, 63–74.

11. Essén H. Darwin magnetic interaction energy and its macroscopic consequences. *Phys. Rev. E*, 1996, v. 53, 5228–5239.

12. Essén H. Phase-space energy of charged particles with negligible radiation: Proof of spontaneous formation of magnetic structures and new effective forces. *Phys. Rev. E*, 1997, v. 56, 5858–5865.

13. Essén H. Magnetism of matter and phase-space energy of charged particle systems. *J. Phys. A: Math. Gen.*, 1999, v. 32, 2297–2314.

14. Kaufman A. N. and Rostler P. S. The Darwin model as a tool for electromagnetic plasma simulation. *Phys. Fluids*, 1971, v. 14, 446–448.

15. Stettner R. Conserved quantities and radiation effects for a closed system of charged particles. *Ann. Phys. (N.Y.)*, 1971, v. 67, 238–251.

16. Boyer T. H. Examples and comments related to relativity controversies. *Am. J. Phys.*, 2012, v. 80, 962–971.

17. Boyer T. H. Faraday induction and the current carriers in a circuit. *Am. J. Phys.*, 2015, v. 83, 263–442, 2015.

18. Krause T. B., Apte A., and Morrison P. J. A unified approach to the Darwin approximation. *Phys. of Plasmas*, 2007, v. 14, 102112–1–10.

19. Essén H. Magnetic dynamics of simple collective modes in a two-sphere plasma model. *Phys. of Plasmas*, 2005, v. 12, 122101–1–7.

20. Essén H. Electrodynamic model connecting superconductor response to magnetic field and to rotation. *Eur. J. Phys.*, 2005, v. 26, 279–285.

21. Essén H. From least action in electrodynamics to magnetomechanical energy – a review. *Eur. J. Phys.*, 2009, v. 30, 515–539.

22. Essén H. Classical diamagnetism, magnetic interaction energies, and repulsive forces in magnetized plasmas. *EPL*, 2011, v. 94, 47003–1–5.

23. Essén H. An exact formula for the electromagnetic momentum in terms of the charge density and the coulomb gauge vector potential. *Eur. J. Phys.*, 2018, v. 39, 025202–1–9.

24. Essén H. and Nordmark A. B. Hamiltonian of a homogeneous two-component plasma. *Phys. Rev. E*, 2004, v. 69, 036404–1–9.

25. Essén H. and Fiolhais M. C. N. Meissner effect, diamagnetism, and classical physics —– a review. *Am. J. Phys.*, 2012, v. 80, 164–169.

26. Woodcock H. W. and Havas P. Approximately relativistic Lagrangians for classical interacting point particles. *Phys. Rev. D*, 1972, v. 12, 3422–3444.

27. Essén H. and Nordmark A. B. unpublished, 2019.

28. Wigner E. On the interaction of electrons in metals. *Phys. Rev.*, 1934, v. 46, 1002–1011.

29. Pereg-Barnea T. and Franz M. Duality and the vibrational modes of a Cooper-pair Wigner crystal. *Phys. Rev. B*, 2006, v. 74, 014518.

30. Dai J. X., Tao W., Hor P., and Dai X. X. A possible mechanism of superconductivity based on Wigner crystal and BEC. *Int. J. Mod. Phys. B*, 1999, v. 13, 3499–3504.

31. Vasiliev B. V. The superconductivity as a consequence of an ordering of the electron gas zero-point oscillations. *Physica C: Superconductivity*, 2011, v. 471, 277–284.

32. Vasiliev B. V. Superconductivity and condensation of ordered zero-point oscillations. *Physica C: Superconductivity*, 2012, v. 483, 233–246.

33. Frenkel J. On a possible explanation of superconductivity. *Phys. Rev.*, 1933, v. 43, 907–912.

34. Essén H. A study of lattice and magnetic interactions of conduction electrons. *Phys. Scr.*, 1995, v. 52, 388–394.

35. Poole C. P. Jr., Prozorov R., Farach H. A., and Creswick R. J. Superconductivity, 3rd edition. Elsevier, Amsterdam, 2014.

36. Kulsrud R. M. Plasma Physics for Astrophysics. Princeton University Press, New Jersey, 2005.

37. Darwin C. G. Inertia of electrons in metals. *Proc. R. Soc. Lond. (UK)*, 1936, v. 154A, 61–66.

38. Assis A.Ḱ. T. and Hernandes J. A. Magnetic energy and effective inertial mass of the conduction electrons in circuit theory. *Electromagnetic Phenomena*, 2006, v. 6, 29–35.

39. Nicastro F., Kaastra J., Krongold Y., Borgani S., Branchini E., Cen R., Dadina M., Danforth C. W., Elvis M., Fiore F., Gupta A., Mathur S., Mayya D., Paerels F., Piro L., Rosa-Gonzalez D., Schaye J., Shull J. M., Torres-Zafra J., Wijers N., and Zappacosta L. Observations of the missing baryons in the warm–hot intergalactic medium. *Nature*, 2018, v. 558, 406–409.

———————————

Quantum State Tomography for Qutrits Subject to Laser Cooling

Artur Czerwinski

Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Grudziadzka 5, 87-100 Torun, Poland.
E-mail: aczerwin@umk.pl

In this article we propose a dynamic quantum state tomography model for qutrits subject to laser cooling. We prove that one can reduce the number of distinct measurement setups required for state reconstruction by employing the stroboscopic approach. The results are in line with current advances in quantum tomography where there is a strong tendency to investigate the optimal criteria for state reconstruction. We believe that the stroboscopic approach can be considered an efficient tool for density matrix identification since it allows to determine the minimal number of distinct observables needed for quantum state tomography.

1 Introduction

The term *quantum tomography* is used in reference to a wide variety of methods which aim to reconstruct the accurate representation of a quantum system by performing a series of measurements. Mathematically, the complete knowledge about the state of a quantum system can be encoded in, for example, the density operator, the wavefunction or the Wigner function. In this article we discuss the problem of the density matrix reconstruction.

One of the most fundamental approaches to quantum state tomography, the so-called static tomography model, enables to reconstruct the density matrix of a quantum system provided one can measure $N^2 - 1$ distinct observables (where $N = \dim \mathcal{H}$). Any density matrix can be decomposed in the basis of SU(N) generators in such a way that the coefficients correspond to the mean values of the operators [1]. This approach has been excessively studied in many papers and books, such as [2, 3]. However, there is a significant disadvantage connected with this method. In a laboratory one usually is not able to define $N^2 - 1$ distinct physical quantities that could be measured.

The most important property that all tomography models should possess is practicability, which means that a theoretical model should have a potential to be implemented in an experiment in the future. Therefore, when dealing with quantum state tomography we should bear in mind the limitations related to laboratory reality. For this reason, in this article we employ the stroboscopic approach to quantum tomography, which for the first time was proposed by Andrzej Jamiolkowski in [4]. Later it was developed in other research papers such as [5] and [6]. In order to get a broad perspective one may also refer to a very well-written review paper [7]. Recently some new results concerning the stroboscopic approach has been presented in [8, 9].

The stroboscopic tomography concentrates on determining the optimal criteria for quantum tomography of open systems. The main goal of this method is to reduce the number of distinct observables required for quantum tomography by utilizing knowledge about time evolution of the system. The data for the density matrix reconstruction is provided by mean values of some hermitian operators $\{Q_1, \ldots, Q_r\}$, where naturally $Q_i = Q_i^*$. The set of observables is not informationally complete, which means that a single measurement of each operator does not provide sufficient information for quantum state reconstruction.

The underlying principle behind the stroboscopic approach claims that if one has the knowledge about the evolution of the system, each observable can be measured repeatedly at a certain number of time instants. Naturally, each individual measurement is performed over a distinct copy of the system since we do not consider the collapse of the quantum state caused by measurements. Therefore, we assume that our source can prepare a large sample of systems in the identical (but unknown) quantum state.

In the stroboscopic approach to quantum tomography the fundamental question that we are interested in concerns the minimal number of distinct observables required for quantum state reconstruction. One can recall the theorem concerning the minimal number of observables [5].

Theorem 1. *For a quantum system with dynamics given by a master equation of the form [10, 11]:*

$$\dot{\rho}(t) = \mathbb{L}[\rho(t)], \tag{1}$$

one can calculate the minimal number of distinct observables for quantum tomography from the formula:

$$\eta := \max_{\lambda \in \sigma(\mathbb{L})} \{\dim \operatorname{Ker}(\mathbb{L} - \lambda \mathbb{I})\}, \tag{2}$$

where by $\sigma(\mathbb{L})$ one should understand the spectrum of the operator $\mathbb{L}$.

The linear operator $\mathbb{L}$ that appears in (1) shall be called the generator of evolution. The number η is usually referred to as *the index of cyclicity* of a quantum system.

The theorem 1 means that for any linear generator $\mathbb{L}$ there exists a set of observables $\{Q_1, \ldots, Q_\eta\}$ such that their ex-

pectation values determine the initial density matrix. Consequently, they also determine the complete trajectory of the state (one can compute the density matrix at any time instant).

If we denote the number of required measurements of each observable from the set $\{Q_1, ..., Q_\eta\}$ by M_i for $i = 1, \ldots, \eta$, then one can also recall the theorem on the upper limit of moments of measurement [6].

Theorem 2. *In order to provide sufficient data for the density matrix reconstruction the number of times that each observable from the set $\{Q_1, ..., Q_\eta\}$ should be measured satisfies the inequality:*

$$M_i \leq \deg \mu(\mathbb{L}), \tag{3}$$

where by $\mu(\mathbb{L})$ we denote the minimal polynomial of $\mathbb{L}$.

The theorem 2 gives the upper boundary concerning the number of measurements of each single observable. One can notice that the ability to compute the minimal polynomial of the generator $\mathbb{L}$ is crucial in order to determine the upper limit for the number of measurements. Naturally, another problem relates to the choice of the time instants. Some considerations about this issue can be found in [6].

In the next section the theorems concerning the stroboscopic tomography shall be applied to three-level quantum systems with the evolution known as laser cooling. This article brings substantial advancement to the field of quantum state tomography. In [8] the author introduced optimal criteria for quantum tomography of qubits. In the current work we proceed towards higher dimensional Hilbert space. We prove that the stroboscopic tomography can be an effective method of state reconstruction for qutrits provided one knows how the system evolves.

2 Quantum tomography schemes for three-level systems subject to laser cooling

2.1 Static approach to quantum tomography of qutrits

In case of three-level quantum systems one would naturally employ the Gell-Mann matrices in order to decompose any density matrix. We follow the original notation from [12] and therefore, the Gell-Mann matrices shall be denoted by $\{\lambda_1, \lambda_2, \ldots, \lambda_8\}$. They have the following forms:

$$\lambda_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \lambda_2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \lambda_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\lambda_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \quad \lambda_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$\lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

The Gell-Mann matrices are the generators of the SU(3) group. They are the generalization of the Pauli operators for three-level systems. They have some algebraic properties which are useful for quantum state tomography, i.e.:

$$\lambda_i = \lambda_i^*, \qquad \operatorname{Tr} \lambda_i = 0 \quad \text{and} \quad \operatorname{Tr} \lambda_i \lambda_j = 2\delta_{ij}. \tag{4}$$

For three-level quantum systems the initial density matrix $\rho(0) \in \mathcal{S}(\mathcal{H})$ can be decomposed in the basis of the Gell-Mann matrices [1]:

$$\rho(0) = \frac{1}{3} \mathbb{I}_3 + \frac{1}{2} \sum_{i=1}^{8} \langle \lambda_i \rangle \lambda_i, \tag{5}$$

where $\langle \lambda_i \rangle$ is the expectation value of the observable λ_i. Mathematically, it can be computed as $\langle \lambda_i \rangle = \operatorname{Tr}\{\lambda_i \rho(0)\}$.

If one would like to directly apply this decomposition in order to reconstruct the density matrix, one would have to know the mean values of eight distinct observables $\{\lambda_1, \lambda_2, \ldots, \lambda_8\}$. Such data would be necessary to complete the formula for $\rho(0)$. This approach to quantum tomography, which does not take advantage of the knowledge about evolution, shall be referred to as the static approach. This scheme appears impractical since one is not able to define eight distinct physical quantities. This observation justifies the need for more economic approach which aims to decrease the number of distinct observables.

2.2 Dynamic approach to quantum state tomography of qutrits

Laser cooling is a very widely investigated topic in modern Physics, e.g. [13, 14]. A lot of attention has been paid to different aspects of this problem. In particular, one may refer to applications of atoms subject to laser cooling in quantum information encoding [15]. In this paper we search for a link between laser cooling and quantum state tomography.

An example often studied in the area of laser spectroscopy is a quantum system subject to laser cooling with three energy levels (dim $\mathcal{H} = 3$) [16]. The evolution of the density matrix of such a three-level system is given by a master equation of the form:

$$\frac{d\rho(t)}{dt} = -i[H(t), \rho(t)] +$$

$$+ \gamma_1 \left(E_1 \rho(t) E_1^* - \frac{1}{2} \{E_1^* E_1, \rho(t)\} \right) + \tag{6}$$

$$+ \gamma_2 \left(E_2 \rho(t) E_2^* - \frac{1}{2} \{E_2^* E_2, \rho(t)\} \right),$$

where $E_1 = |1\rangle \langle 2|$ and $E_2 = |3\rangle \langle 2|$. The vectors $\{|1\rangle, |2\rangle, |3\rangle\}$ denote the standard basis in $\mathcal{H}$.

This kind of dynamics appears when the excited state $|2\rangle$ decays spontaneously into two ground states $|1\rangle$ and $|3\rangle$ with corresponding decoherence rates γ_1 and γ_3.

Moreover in this analysis we take $H(t) = [0]$, where $[0]$ denotes a 3−dimensional matrix with all entries equal 0. This assumption means that we shall analyze only the Lindbladian part of the evolution equation.

In case of a three-level open quantum system with dynamics given by the master equation from (6) we can formulate and prove a theorem which provides the minimal number of distinct observables required for quantum tomography.

Theorem 3. *For a quantum system subject to laser cooling according to (6) there exists four distinct observables such that their average values (measured at selected time instants over different copies of the system) suffice to determine the initial density matrix $\rho(0)$.*

Proof. Based on the method of matrix vectorization [8, 17], the dissipative part of the generator of evolution (6) can be explicitly expressed as a matrix:

$$\mathbb{L} = \gamma_1 \left(E_1 \otimes E_1 - \frac{1}{2} \left(\mathbb{I}_9 \otimes E_1^T E_1 + E_1^T E_1 \otimes \mathbb{I}_9 \right) \right) + $$
$$+ \gamma_2 \left(E_2 \otimes E_2 - \frac{1}{2} \left(\mathbb{I}_9 \otimes E_2^T E_2 + E_2^T E_2 \otimes \mathbb{I}_9 \right) \right). \quad (7)$$

Taking into account the fact that the vectors $\{|1\rangle, |2\rangle, |3\rangle\}$ constitute the standard basis, the matrix form of the quantum generator $\mathbb{L}$ can be obtained:

$$\mathbb{L} = \begin{bmatrix} 0 & 0 & 0 & 0 & \gamma_1 & 0 & 0 & 0 & 0 \\ 0 & -\Gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\Gamma & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2\Gamma & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\Gamma & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\Gamma & 0 \\ 0 & 0 & 0 & 0 & \gamma_2 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (8)$$

where $\Gamma = \frac{1}{2}(\gamma_1 + \gamma_2)$.

Having the matrix form of the generator of evolution $\mathbb{L}$, one can calculate its eigenvalues:

$$\sigma(\mathbb{L}) = \{0, 0, 0, 0, -2\Gamma, -\Gamma, -\Gamma, -\Gamma, -\Gamma\}. \quad (9)$$

Since in this case the operator $\mathbb{L}$ is not self-adjoint, the algebraic multiplicity of an eigenvalue does not have to be equal to its geometric multiplicity. But one can quickly determine that there are four linearly independent eigenvectors that correspond to the eigenvalue 0. Therefore, we can find the index of cyclicity for the operator in question:

$$\eta = \max_{\lambda \in \sigma(\mathbb{L})} \{\dim \mathrm{Ker}(\mathbb{L} - \lambda \mathbb{I}_9)\} = 4, \quad (10)$$

which means that we need exactly four distinct observables to perform quantum tomography of the analyzed system. □

One can instantly notice that if the static approach was applied to three-level laser cooling, one would have to measure 8 distinct observables whereas in the dynamic approach 4 observables suffice to perform quantum tomography. If one thinks of potential applications in experiments, then our result means that one would have to prepare 4 different experimental setups instead of 8. This observation demonstrates that the stroboscopic approach has an advantage over the static approach because it is more economic when it comes to the number of distinct kinds of measurement.

The next issue that we are interested in is the minimal polynomial for the operator $\mathbb{L}$. Assuming that this polynomial has the monic form, i.e.:

$$d_3 \mathbb{L}^3 + d_2 \mathbb{L}^2 + d_1 \mathbb{L} + d_0 \mathbb{I} = 0, \quad (11)$$

one can get :

$$d_3 = 1, \ d_2 = \frac{3}{2}(\gamma_1 + \gamma_2), \ d_1 = \frac{1}{2}(\gamma_1 + \gamma_2)^2, \ d_0 = 0. \quad (12)$$

Thus, we see that $\deg \mu(\mathbb{L}) = 3$. This means that each observable should be measured at most at three different time instants. One can conclude that, since we need 8 independent pieces of information to reconstruct the initial density matrix, not every observable will be measured the maximum number of times. To provide a precise answer to the question concerning the algebraic structure of the observables and the choice of time instants, we shall accept additional assumptions concerning the generator of evolution.

Let us consider a special case of the generator of evolution defined in (8) such that $\gamma_1 = 1/4$ and $\gamma_2 = 3/4$. For this specific generator, we can formulate a theorem.

Theorem 4. *The initial density matrix $\rho(0)$ of a three-level system subject to laser cooling can be reconstructed from the mean values of four observables of the form:*

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1+i \\ 0 & 1-i & 0 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0 & 0 & 1+i \\ 0 & 0 & 0 \\ 1-i & 0 & 0 \end{bmatrix},$$

$$Q_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & -\frac{2}{\sqrt{3}} \end{bmatrix}, \quad Q_4 = \begin{bmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (13)$$

where the mean values of Q_1 and Q_2 are measured at 3 distinct time instants and the observables Q_3 and Q_4 once at $t = 0$.

Proof. According to the assumptions of the stroboscopic tomography, the information that one can obtain from an experiment is encoded in the mean values of some observables, which mathematically can be written as:

$$m_i(t_j) = \mathrm{Tr}\{Q_i \rho(t_j)\}, \quad (14)$$

where $\rho(t_j) = \exp(\mathbb{L} t_j)[\rho(0)]$.

One is aware that $\exp(\mathbb{L}t_j)$ can be decomposed as:

$$\exp(\mathbb{L}t) = \alpha_0(t)\mathbb{I}_9 + \alpha_1(t)\mathbb{L} + \alpha_2(t)\mathbb{L}^2, \qquad (15)$$

where the functions $\{\alpha_0(t), \alpha_1(t), \alpha_2(t)\}$ are linearly independent. In order to determine these functions, we need to employ the minimal polynomial of $\mathbb{L}$ and then solve a system of differential equations [6, 8]. Having done the necessary computations, one gets:

$$\begin{cases} \alpha_0(t) = 1 \\ \alpha_1(t) = e^{-t} - 4e^{-\frac{1}{2}t} + 3 \\ \alpha_2(t) = 2e^{-t} - 4e^{-\frac{1}{2}t} + 2. \end{cases} \qquad (16)$$

Since one is able to decompose $\exp(\mathbb{L}t_j)$ in the basis of three operators $\{\mathbb{I}_9, \mathbb{L}, \mathbb{L}^2\}$ due to linearity of the matrix trace we get:

$$m_i(t_j) = \alpha_0(t_j)\mathrm{Tr}\{Q_i\rho(0)\} + \alpha_1(t_j)\mathrm{Tr}\{Q_i\mathbb{L}[\rho(0)]\} +$$
$$+ \alpha_2(t_j)\mathrm{Tr}\{Q_i\mathbb{L}^2[\rho(0)]\}. \qquad (17)$$

If by $\mathbb{L}^*$ we shall denote the dual operator to $\mathbb{L}$, then by changing the perspective from the Schrödinger picture to the Heisenberg representation we can obtain:

$$m_i(t_j) = \alpha_0(t_j)\mathrm{Tr}\{Q_i\rho(0)\} + \alpha_1(t_j)\mathrm{Tr}\{\mathbb{L}^*[Q_i]\rho(0)\} +$$
$$+ \alpha_2(t_j)\mathrm{Tr}\{(\mathbb{L}^*)^2[Q_i]\rho(0)\}. \qquad (18)$$

This means that if the mean value of the observable Q_1 is measured at three distinct time instants, one gets a matrix equation:

$$\begin{bmatrix} m_1(t_1) \\ m_1(t_2) \\ m_1(t_3) \end{bmatrix} = \begin{bmatrix} \alpha_0(t_1) & \alpha_1(t_1) & \alpha_2(t_1) \\ \alpha_0(t_2) & \alpha_1(t_2) & \alpha_2(t_2) \\ \alpha_0(t_3) & \alpha_1(t_3) & \alpha_2(t_3) \end{bmatrix} \begin{bmatrix} \mathrm{Tr}\{Q_1\rho(0)\} \\ \mathrm{Tr}\{\mathbb{L}^*[Q_1]\rho(0)\} \\ \mathrm{Tr}\{(\mathbb{L}^*)^2[Q_1]\rho(0)\} \end{bmatrix} \quad (19)$$

Since the functions $\{\alpha_0(t), \alpha_1(t), \alpha_2(t)\}$ are linearly independent one can agree that if we select three different non-zero time instants such that $t_1 \neq t_2 \neq t_3$, then the matrix $[\alpha_k(t_j)]$ must be invertible. It implies that the measurement results $\{m_1(t_1), m_1(t_2), m_1(t_3)\}$ can be translated into a set of scalar products:

$$\left\{ \mathrm{Tr}\{Q_1\rho(0)\}, \mathrm{Tr}\{\mathbb{L}^*[Q_1]\rho(0)\}, \mathrm{Tr}\{(\mathbb{L}^*)^2[Q_1]\rho(0)\} \right\}.$$

The very same measurement procedure, which must result in a matrix equation analogous to (19), can be performed for the observable Q_2. Triple measurement of Q_2 at distinct time instants yields a set of the scalar products:

$$\left\{ \mathrm{Tr}\{Q_2\rho(0)\}, \mathrm{Tr}\{\mathbb{L}^*[Q_2]\rho(0)\}, \mathrm{Tr}\{(\mathbb{L}^*)^2[Q_2]\rho(0)\} \right\}.$$

Finally, a single measurement of the average value of Q_3 and Q_4 at time instant $t = 0$ provides another two scalar products: $\{\mathrm{Tr}\{Q_3\rho(0)\}, \mathrm{Tr}\{Q_4\rho(0)\}\}$.

One can check numerically that the operators:

$$\left\{ \mathbb{I}_3, Q_1, \mathbb{L}^*[Q_1], (\mathbb{L}^*)^2[Q_1], Q_2, \mathbb{L}^*[Q_2], (\mathbb{L}^*)^2[Q_2], Q_3, Q_4 \right\}$$

constitute a spanning set (they are all linearly independent), which means that they span the space $B_*(\mathcal{H})$.

The spanning criterion is the necessary and sufficient condition for the ability to reconstruct the initial density matrix of a qutrit subject to laser cooling. This condition is satisfied for the observables defined in the theorem 4, which can be observed numerically by using the software Mathematica 11.

In other words, the operators:

$$\left\{ \mathbb{I}_3, Q_1, \mathbb{L}^*[Q_1], (\mathbb{L}^*)^2[Q_1], Q_2, \mathbb{L}^*[Q_2], (\mathbb{L}^*)^2[Q_2], Q_3, Q_4 \right\}$$

constitute a quorum, i.e. they span the space to which $\rho(0)$ belongs. Therefore, the scalar products that one can calculate from the measurement results can be considered a complete set of information. Thus, the measurement procedure, which utilizes only 4 distinct kinds of measurement, provides 8 independent pieces of information which are sufficient for the density matrix reconstruction.   □

The theorems 3 and 4 provide a complete description of the quantum tomography scheme. One knows exactly what steps should be taken in order to compute the unknown density matrix.

The results are in accord with current trends in quantum state tomography where a lot of attention is paid to the methods which aim to reduce the experimental effort, e.g. [18, 19]. If one can access the knowledge about dynamics of the system encoded in the generator of evolution, it seems more convenient to perform repeatedly the same kind of measurement (over distinct copies of the system) rather than develop a large number of different experimental setups.

3 Summary

In this paper we presented a complete quantum tomography model for qutrits subject to laser cooling. The stroboscopic approach was applied to determine the optimal criteria for density matrix reconstruction. It was demonstrated that one can reduce the number of distinct observables by 50% provided the knowledge about evolution is applied. The algebraic structure of the observables was presented along with a detailed description of the scheme. Dynamic methods of state reconstruction appear to be very practical since they allow to retrieve the initial density matrix in the most economical way, by minimizing the number of distinct measurement setups.

The article indicates a link between quantum state tomography and laser cooling. Both topics play a substantial role in the field of quantum communication. The ability to reconstruct the quantum state from measurements is crucial to determine the efficiency of quantum communication protocols. Whereas atoms subject to laser cooling are often utilized to

encode quantum information. The dynamic quantum tomography scheme presented in this article combines these two lines of research.

The current work can be extended in the future research by studying the problem of quantum state tomography for systems subject to laser cooling with more than three energy levels. This task requires advanced algebraic methods to study the spectrum of the generator of evolution as well as to determine its minimal polynomial.

Received on March 4, 2020

References

1. Kimura G. The Bloch vector for N-level systems. *Phys. Lett. A*, 2003, v. 314 (5-6), 339–349.

2. Altepeter J. B., James D. F. V., Kwiat P. G. 4 Qubit Quantum State Tomography. In: Paris M. G. A., Rehacek J. (eds.) Quantum State Estimation. Springer, Berlin, 2004, 111–145.

3. Alicki R., Lendi K. Quantum Dynamical Semigroups and Applications. Springer, Berlin, 1987.

4. Jamiolkowski A. The Minimal Number of Operators for Observability of N-level Quantum Systems. *Int. J. Theor. Phys.*, 1983, v. 22 (4), 369–376.

5. Jamiolkowski A. On complete and incomplete sets of observables, the principle of maximum entropy - revisited. *Rep. Math. Phys.*, 2000, v. 46 (3), 469–482.

6. Jamiolkowski A. On a Stroboscopic Approach to Quantum Tomography of Qudits Governed by Gaussian Semigroups. *Open Syst. Inf. Dyn.*, 2004, v. 11 (1), 63–70.

7. Jamiolkowski A. Fusion Frames and Dynamics of Open Quantum Systems. In: Lyagushyn S. (ed.) Quantum Optics and Laser Experiments. InTech, Rijeka, 2012, 67–84.

8. Czerwinski A. Applications of the Stroboscopic Tomography to Selected 2-Level Decoherence Models. *Int. J. Theor. Phys.*, 2016, v. 55 (2), 658–668.

9. Czerwinski A. Minimal number of observables for quantum tomography of systems with evolution given by double commutators. *Quantum Stud.: Math. Found.*, 2017, v. 4 (4), 287–294.

10. Gorini V., Kossakowski A., Sudarshan E. C. G. Completely positive dynamical semigroups of n-level systems. *J. Math. Phys.*, 1976, v. 17 (5), 821–825.

11. Lindblad G. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.*, 1976, v. 48 (2), 119–130.

12. Gell-Mann M. Symmetries of Baryons and Mesons. *Phys. Rev.*, 1962, v. 125 (3), 1067–1084.

13. Bartana A., Kosloff R., Tannor D. J. Laser cooling of molecular internal degrees of freedom by a series of shaped pulses. *J. Chem. Phys.*, 1993, v. 99 (1), 196–210.

14. Tannor D. J., Bartana A. On the Interplay of Control Fields and Spontaneous Emission in Laser Cooling. *J. Phys. Chem. A*, 1999, v. 103 (49), 10359–10363.

15. Saffman M., Walker T. G., Molmer K. Quantum information with Rydberg atoms. *Rev. Mod. Phys.*, 2010, v. 82 (3), 2313.

16. Sklarz S. E., Tannor D. J., Khaneja N. Optimal Control of Quantum Dissipative Dynamics: Analytic solution for cooling the three level Λ system. *Phys. Rev. A*, 2004, v. 69 (5), 053408.

17. Henderson H. V., Searle S. R. The vec-permutation matrix, the vec operator and Kronecker products: A review. *Linear and Multilinear A.*, 1981, v. 9 (4), 271–288.

18. Czerwinski A. Optimal evolution models for quantum tomography. *J. Phys. A: Math. Theor.*, 2016, v. 49 (7), 075301.

19. Oren D. *et al.* Quantum state tomography with a single measurement setup. *Optica*, 2017, v. 4 (8), 993–999.

————————

Can the Nuclear Liquid Drop Model Be Improved?

Omar Yépez

E-mail: yepezoj@gmail.com

To be part of a nucleus, the constituent nucleons lose part of the original area they have. This can be measured by subtracting this area from the surface area of the nucleus. This was measured and plotted against the respective nuclear binding energy. A straight linear relationship was found for all elements, light or heavy. For a given element, the nuclear binding energy is inversely proportional to the lost original area. Thus meaning, that more area lost corresponded to a larger binding energy. β^- decay occurred to produce a nucleus with less loss of the nucleons' original area. β^+ decay occurred to produce a nucleus with less Coulomb repulsion. The nucleus stability just follows a trade-off between these two trends.

1 Introduction

Even though there is a very complete understanding of nuclear forces, they are so complicated that this knowledge can not be used to construct a complete theory of the nucleus. In other words, it is not possible to explain all nuclei properties based on the nuclear force acting between protons and neutrons. However, there is a number of models, or rudimentary theories with certain validity, which can explain a limited number of certain properties. In between those theories, the liquid drop model has been used with success and it has not changed for more than sixty years [1]. Theoretically, the nuclear liquid drop model calculates the nuclear binding energy by taking into account a number of interactions [2], i.e.

$$E_b = a_V A - a_S A^{2/3} - a_C \frac{Z(Z-1)}{A^{1/3}} - a_A \frac{(A-2Z)^2}{A} \pm \delta(A,Z) \quad (1)$$

where the coefficients a_V, a_S, a_C, a_A and $\delta(A,Z)$ are determined empirically. The volume of the nucleus is proportional to A, thus the term $a_V A$. Nucleons on the surface of the nucleus have fewer nearest neighbors. This can also be thought of as a surface tension term. If the volume term is proportional to A, the surface term should be proportional to $A^{1/3}$. The Coulomb term is due to the electric repulsion between protons in the nucleus. The asymmetry term a_A is due to the Pauli exclusion principle and the pairing term which capture the effect of spin-coupling. This formula gives the nuclear binding energy with a positive sign for exothermic reactions.

Besides its original success and continuous efforts, this model has not progressed more and still does not perform well with light nuclei [1]. There could be a number of reasons for that. Forcing a correlation between the nuclear binding energy against the number of nucleons, A; or putting several parameters to be fit against powers of A could be some of the reasons.

Nowadays, there is plenty of data about the radiuses of all isotopes for all elements, which are reported in [3]. Thus, a better correlation between the nuclear binding energy and the nucleons' surface term could be achieved. In this paper, a straight linear correlation was found between a geometrical construct that measures how much surface area has been lost by a given isotope's nucleons (Ω) and its nuclear binding energy. Changes between parent and daughter nucleus' Ω and the Coulomb repulsion are sufficient to explain β decay, emission of protons, α particles and neutrons, as well as electron capture. The nucleus stability appears as a consequence of a trade-off between these two trends.

2 Experimental

All isotope radiuses were reported in [3]. The radiuses of the proton and neutron used were: $r_p = 0.8783$ fm [3] and $r_n = 1.21$ fm [4], respectively. Assuming they are all spheres*, the formula created to compute how much of the nucleons spherical surface area has been lost or gained to form the nucleus was

$$\Omega = \frac{4\pi(r_i^2 - Zr_p^2 - Nr_n^2)}{Z+N} . \quad (2)$$

Ω is the surface area difference between the isotope and its components per number of nucleons, $A = Z + N$, in fm^2, r_i is the radius of the isotope, Z is the number of protons and N is the number of neutrons. The nuclear binding energy (mass defect) was calculated by the following formula [5]

$$E_b = (Zm_e + Zm_p + Nm_n - m_i)c^2 \quad (3)$$

where m_e, m_p and m_n are the masses of the electron, proton and the neutron respectively and m_i is the mass of the isotope. The masses of the isotopes were reported in [6], the decay mode, energy and yields were reported in [7]. The following figures present the graphs of Ω versus the nuclear binding energy for different elements. In the case of nuclear decays, $\Delta\Omega$ is the difference between daughter and parent nucleus' Ω.

3 Results

Fig. 1 shows that Ω for a given group of isotopes is inversely proportional to its nuclear binding energy. It is also observed that the rate of its change diminished as the number of protons increase. In this way, helium presents the largest changes in

---

*It is known the nucleus has different shapes. A sphere is one of them.

Ω within smaller changes in nuclear binding energy, whereas radon showed very small changes in Ω corresponding to larger changes in binding energy.

Fig. 2 presents Ω versus nuclear binding energy for He, Li, Be and B isotopes. The isotope with a red circle are the stable ones. It is clearly observed that as the binding energy increases, the nucleons of a given isotope presents a more negative Ω and requires more binding energy to form.

Beginning with two stable isotopes, ^{3}He's Ω is positive because the addition of the area of two protons and one neutron is not larger than the area of the isotope. Whereas ^{4}He's Ω is negative because the addition of the areas of two protons and two neutrons is larger than the area of that isotope. Once ^{6}He formed, the stability is lost. Given that ^{6}Li has a lower mass than ^{6}He, β^- decays occur, liberating 3.51 MeV. This process follows an Ω increase and therefore $\Delta\Omega$ was 6.48 fm^2 for this reaction.

In the same manner, ^{8}He suffers β^- decay and neutron emission to ^{7}Li, with 16% reaction yield. It liberates 8.63 MeV. This is also accompanied by the emission of one neutron. Again, the daughter nucleus presents a more positive Ω and therefore $\Delta\Omega = 6.41$ fm^2 for this reaction.

^{8}He also suffers β^- decay to ^{8}Li, with 83% yield. It liberates 10.66 MeV and $\Delta\Omega = 3.86$ fm^2.

^{7}Be suffers 100% β^+ decay into ^{7}Li. Contrary to the previous trend, in this process the daughter presented a more negative Ω than the parent nucleus. But also, β^+ diminished the number of protons in the daughter nucleus, thus diminishing the Coulomb repulsion. Contrary to previous β^- decay, in this case $\Delta\Omega = -3.13$ fm^2.

^{9}Li repeats ^{6}He's behavior. ^{11}Li presents neutron emission to ^{10}Be with 86.3% yield and β^- decay to ^{11}Be with 6% yield*. This is very similar to ^{8}He transmutation. Finally, ^{10}Be repeats ^{6}He's behavior. Table 1 summarizes the nuclear processes observed in Fig. 2. It is clearly observed that β^- and neutron emission presents a positive $\Delta\Omega$, whereas β^+ decay shows a negative $\Delta\Omega$.

Fig. 3 presents Ω versus nuclear binding energy for O, F, Ne, Na and Mg isotopes. A 100% of ^{17}Ne transmutes to ^{16}O after β^+ decay and a proton emission, producing 11.63 MeV. $\Delta\Omega$ in this case was -1.88 fm^2. A 100% of ^{19}Ne transmutes to ^{19}F after β^+ decay, producing 2.20 MeV and $\Delta\Omega = -0.88$ fm^2. ^{20}Na goes to ^{20}Ne with 75% yield, producing 12.87 MeV and $\Delta\Omega = -0.26$ fm^2. It also emits an alpha particle and a positron to produce ^{16}O with 25% yield, generating 8.14 MeV and $\Delta\Omega = -0.26$ fm^2. Table 2 presents the transitions observed in Fig. 3. It is clearly observed that β^+, proton and alpha particle emissions present a negative $\Delta\Omega$, whereas β^- and 2β^- decays show a positive $\Delta\Omega$.

Fig. 4 presents Ω versus nuclear binding energy for Ar, K, Ca, Sc and Ti isotopes. A 100% of ^{38}K transmutes to ^{38}Ar

---

*This nucleus also experiences double and triple neutron emission, α emission and fission in lower yields.
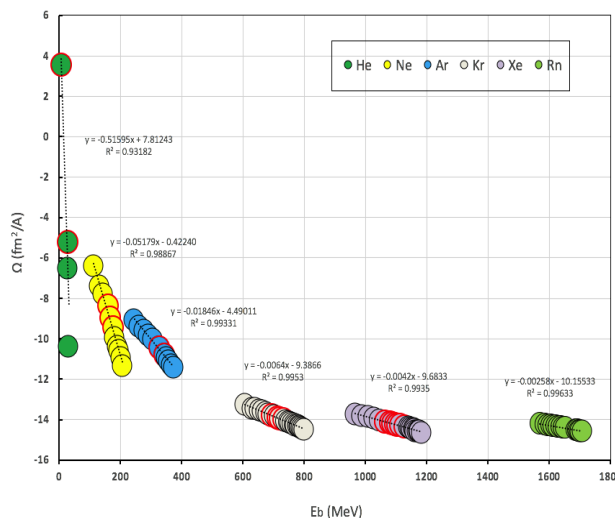
Fig. 1: Ω vs. binding energy for Noble gases. The red circles are the stable isotopes.
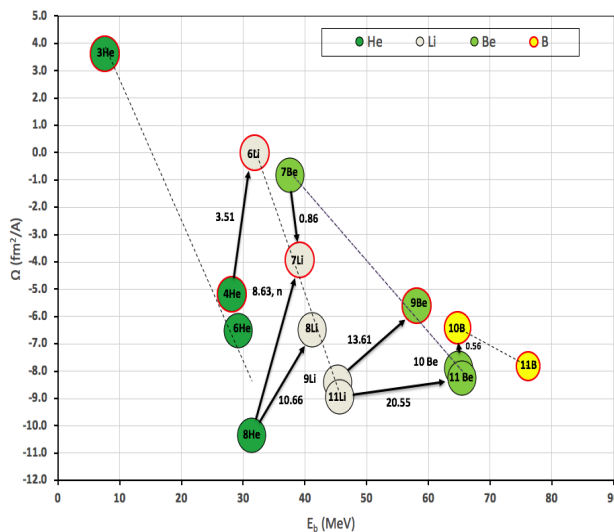
Fig. 2: Ω vs. mass defect for He, Li, Be and B isotopes. The red circles are the stable isotopes. The energy of the transitions (MeV) were reported in [7].

after β^+ decay, producing 4.89 MeV and $\Delta\Omega$ in this case was -0.28 fm^2.

A 100% of ^{39}Ca transmutes to ^{39}K after β^+ decay, producing 6.52 MeV and $\Delta\Omega = -0.28$ fm^2. ^{40}K goes to ^{40}Ca with 89.28% yield, producing 1.31 MeV and $\Delta\Omega = 0.30$ fm^2. ^{40}K also suffers electron capture to ^{40}Ar with 10.72% yield, producing 0.48 MeV and $\Delta\Omega = -0.24$ fm^2. A 100% of ^{41}Ca transmutes to ^{41}K after β^+ decay, producing 0.42 MeV and $\Delta\Omega = -0.32$ fm^2. Also, ^{41}Ar suffers β^- decay to ^{41}Ca producing 2.49 MeV and $\Delta\Omega = 0.21$ fm^2. Table 3 depicts the transitions

Table 1: Reaction, mass and area (ΔΩ) difference between parent and daughter nuclei, and decay mode for the reactions depicted in Figure 2.

Reaction	Released Energy (MeV) [7]	ΔΩ (fm²)	Decay
$^{4}_{2}\text{He} \rightarrow\ ^{6}_{3}\text{Li} + e^{-} + \nu$	3.51	6.48	β⁻
$^{8}_{2}\text{He} \rightarrow\ ^{7}_{3}\text{Li} + e^{-} + n + \nu$	8.63	6.41	β⁻ and neutron emission
$^{8}_{2}\text{He} \rightarrow\ ^{8}_{3}\text{Li} + e^{-} + \nu$	10.66	3.86	β⁻
$^{7}_{4}\text{Be} \rightarrow\ ^{7}_{3}\text{Li} + e^{+} + \nu$	0.86	-3.13	β⁺
$^{9}_{3}\text{Li} \rightarrow\ ^{9}_{4}\text{Be} + e^{-} + \nu$	13.61	2.79	β⁻
$^{11}_{3}\text{Li} \rightarrow\ ^{10}_{4}\text{Be} + e^{-} + n + \nu$	20.55	0.68	β⁻ and neutron emission
$^{10}_{4}\text{Be} \rightarrow\ ^{9}_{5}\text{B} + e^{+} + \nu$	0.56	1.85	β⁻

Table 2: Reaction, mass and area (ΔΩ) difference between parent and daughter nuclei, and decay mode for the reactions depicted in Figure 3.

| Reaction	Released Energy (MeV) [7]	ΔΩ (fm²)	Decay
$^{17}_{10}\text{Ne} \rightarrow\ ^{16}_{8}\text{O} + e^{+} + p + \nu$	14.55	-1.88	β⁺ and proton emission
$^{19}_{10}\text{Ne} \rightarrow\ ^{19}_{9}\text{F} + e^{+} + \nu$	3.24	-0.88	β⁺
$^{20}_{11}\text{Na} \rightarrow\ ^{16}_{8}\text{O} + e^{+} + \alpha + \nu$	9.16	-0.26	β⁺ and α emission
$^{20}_{11}\text{Na} \rightarrow\ ^{20}_{10}\text{Ne} + e^{+} + \nu$	13.89	-0.30	β⁺
$^{21}_{11}\text{Na} \rightarrow\ ^{21}_{10}\text{Ne} + e^{+} + \nu$	3.55	-0.57	β⁺
$^{22}_{11}\text{Na} \rightarrow\ ^{22}_{10}\text{Ne} + e^{+} + \nu$	2.84	-0.50	β⁺
$^{23}_{10}\text{Ne} \rightarrow\ ^{23}_{11}\text{Na} + e^{-} + \nu$	4.38	0.65	β⁻
$^{24}_{10}\text{Ne} \rightarrow\ ^{24}_{12}\text{Mg} + 2e^{-} + 2\nu$	7.99	1.21	2β⁻
$^{25}_{10}\text{Ne} \rightarrow\ ^{25}_{12}\text{Mg} + 2e^{-} + 2\nu$	11.15	0.99	2β⁻
$^{26}_{10}\text{Ne} \rightarrow\ ^{26}_{12}\text{Mg} + 2e^{-} + 2\nu$	16.70	0.98	2β⁻
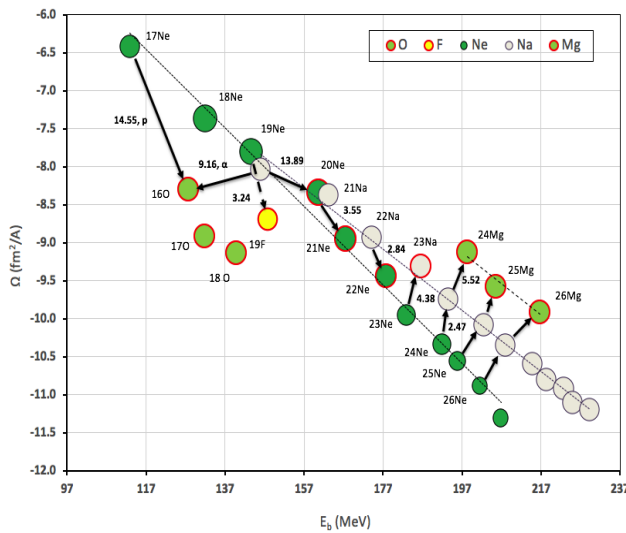
Fig. 3: Ω vs. mass defect for O, F, Ne, Na and Mg isotopes. The red circles are the stable isotopes. The energy of the transitions (MeV) were reported in [7].
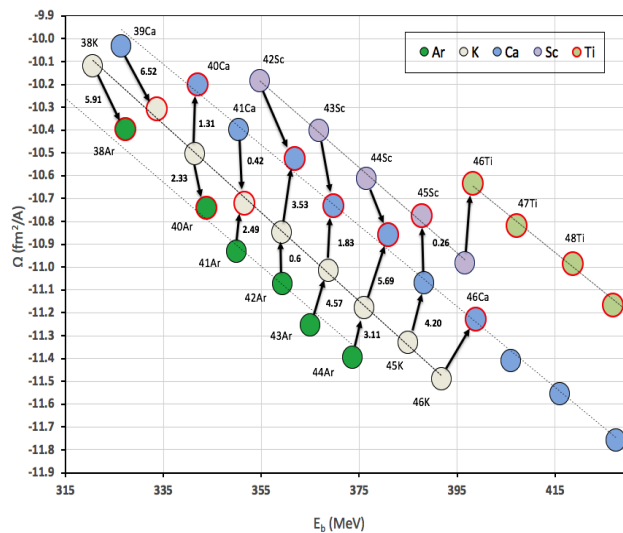
Fig. 4: Ω vs. mass defect for Ar, K, Ca, Sc and Ti isotopes. The red circles are the stable isotopes. The energy of the transitions (MeV) were reported in [7].

observed in Fig. 4. It is clearly observed that electron capture presents a negative ΔΩ.

4 Discussion

4.1 Meaning of Ω and the Nuclear Liquid Drop Model

Ω was computed by using one dimension (the radius) and the three dimensions (the volume). All elements kept a good linear relationship between Ω and the nuclear binding energy. However, in the case of helium, either the linear relationship was lost or the isotopes did not occur proportionally. For example: ^{6}He occurred between ^{3}He and ^{4}He. This relationship is also very sensitive to the neutron radius. Overall, to keep ^{4}He to land between ^{3}He and ^{6}He, r_n needs to be at least 0.05

fm larger than r_p. This may be an indication that the spherical model is only partly applicable to helium. According to the results presented in Fig. 1, it seems that a surface-based Ω is a fundamental property of the isotopes of any element. Given the nature of Ω, it is obvious that larger changes per nucleon would occur in the lowest mass element, helium. This is because the number of nucleons is the lowest. As the number of protons increase, Ω changes less because it is divided by a progressively larger number of nucleons. In a given element, Ω becomes more negative because the addition of the area of the components of the nucleus is progressively larger than its isotope's area. This corresponds to an increasing nuclear binging energy. Which can be interpreted as more energy is

Table 3: Reaction, mass and area (ΔΩ) difference between parent and daughter nuclei, and decay mode for the reactions depicted in Figure 4.

Reaction	Released Energy (MeV) [7]	ΔΩ (fm²)	Decay
${}^{38}_{19}\text{K} \rightarrow {}^{38}_{18}\text{Ar} + e^+ + \nu$	5.91	-0.28	β^+
${}^{39}_{20}\text{Ca} \rightarrow {}^{39}_{19}\text{K} + e^+ + \nu$	6.52	-0.28	β^+
${}^{40}_{19}\text{K} \rightarrow {}^{40}_{18}\text{Ca} + e^- + \nu$	1.31	0.30	β^-
${}^{40}_{19}\text{K} \rightarrow {}^{40}_{18}\text{Ar} + e^+ + \nu$	1.50	-0.24	EC
${}^{41}_{20}\text{Ca} \rightarrow {}^{41}_{19}\text{K} + e^+ + \nu$	0.41	-0.32	β^+
${}^{41}_{18}\text{Ar} \rightarrow {}^{41}_{19}\text{K} + e^- + \nu$	2.49	0.21	β^-
${}^{42}_{18}\text{Ar} \rightarrow {}^{42}_{20}\text{Ca} + 2e^- + 2\nu$	4.10	0.55	$2\beta^-$
${}^{43}_{18}\text{Ar} \rightarrow {}^{43}_{20}\text{Ca} + 2e^- + 2\nu$	6.40	0.53	$2\beta^-$
${}^{44}_{18}\text{Ar} \rightarrow {}^{44}_{20}\text{Ca} + 2e^- + 2\nu$	9.07	0.54	$2\beta^-$
${}^{45}_{19}\text{K} \rightarrow {}^{45}_{21}\text{Sc} + 2e^- + 2\nu$	4.46	0.55	$2\beta^-$

Fig. 5: Ω vs. mass defect for B isotopes. The red circle is the stable isotope.

needed to compress the nucleons' area into the nucleus. This means that all nucleons share the nucleus surface.

This proportionality between the nuclear binding energy and the surface lost to create the nucleus contrasts with the semi-empirical mass formula (1). This is because Fig. 1 presents explicitly that the nuclear binding energy is just proportional to the normalized nucleons' surface area lost to form the isotope. As will be discussed, the other important term is the Coulomb repulsion. This makes (1) to have too many terms to fit. This is because the underlying model for (1) is a sphere-like structure with the neutrons and protons gathered together but still separated as individual spherical particles. The underlying model that Fig. 1 suggests is one where all nucleons share the surface of the nucleus. Which means that protons and neutrons are blended, fused.

4.2 Calculation of ^{8}Be's radius

Not shown in Fig. 2, ^{8}Li transmutes to ^{8}Be and this decays into two ^{4}He. ^{8}Be is not shown in Fig. 2 because its radius was not reported in [3]. An estimation of ^{8}Be's radius can be accomplished by using the inverse proportion between Ω and the other Be isotopes. Fig. 5 shows the result. ^{8}Be nuclear binding energy is 56.50 MeV. Thus, its Ω = -5.65 fm² and the calculated ^{8}Be radius was 2.31 fm. This puts ^{8}Be and ^{9}Be at the same Ω as shown in Fig. 5.

4.3 Why a decay occurs

Fig. 2 depicts the helium isotopes in more detail. Given that ^{2}He is unstable, it seems that helium needs at least one neutron for stability, which occurs in ^{3}He. This suggests the neutron is acting as a Coulomb repulsion insulator. This effect continues in ^{4}He. However, ^{5}He and heavier isotopes become unstable again. It seems that there is a limit to how much area can be lost from the nucleons to form the nucleus, after which a decay is needed to resolve the instability. The first beta decay occurs between the more massive parent ^{6}He and
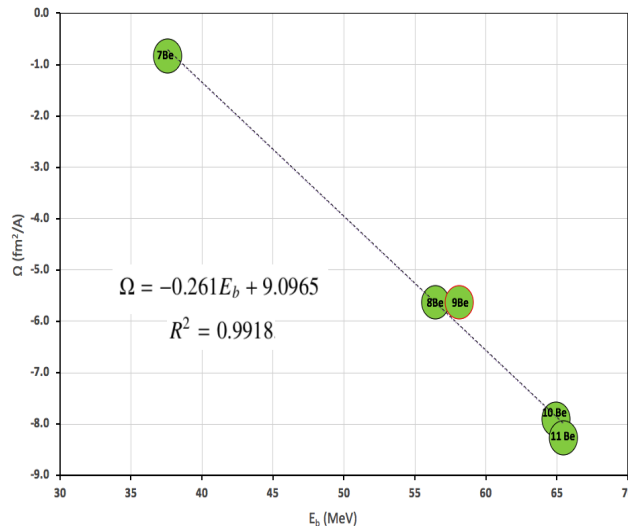
the lighter daughter ^{6}Li producing 3.51 MeV. As observed, β^- decay involves: *to go from a heavier and lower Coulomb repulsion, which has more nucleons' surface area lost (NSL), to a lighter and higher Coulomb repulsion, which has less NSL.* Therefore, the driving force for β^- decay is to reduce the NSL. This is why the $\Delta\Omega$ for this reaction is positive. This is a feature of β^- decay and several examples where $\Delta\Omega$ is positive are shown in Tables 1, 2 and 3. In a more complicated process with 16% reaction yield , ^{8}He suffered neutron emission and β^- decay to transmute to ^{7}Li. This process, nevertheless, has the same features already described for β^- decay, i.e. in neutron emission $\Delta\Omega$ is also positive. Another example of a positive Ω is ^{11}Li going to ^{10}Be.

^{7}Be is the first example of β^+ decay to ^{7}Li. As observed, this process involves: *to go from a heavier and higher Coulomb repulsion nucleus, which has less NSL, to a lighter and lower Coulomb repulsion nucleus, which has more NSL.* This is why the $\Delta\Omega$ for this reaction is negative. Hence, the driving force for β^+ decay is to reduce the Coulomb repulsion. Other examples can be observed in Tables 2 and 3.

Fig. 3 shows that: a) ^{17}Ne transmutes to ^{16}O with 100% yield suffering β^+ decay and proton emission and b) ^{20}Na transforms into ^{16}O by the emission of an α particle and a positron. In both cases, $\Delta\Omega$ is negative. Therefore, these processes are driven by the reduction of Coulomb repulsion.

Fig. 4 presents ^{40}K suffering β^- decay to ^{40}Ca with 89.28% yield. This overwhelms the β^+ decay to ^{40}Ar with 10.72% yield. This reaction suggests that, in this case, to reduce the nucleons' surface area lost is more favorable than to reduce its Coulomb repulsion.

4.4 Nucleus stability

It seems that there is a trade-off between the NSL and Coulomb repulsion for nucleus stability. In Fig. 2, ^{3}He increases the NSL until it reaches ^{6}He. Then, β decay increases the number of protons to produce ^{6}Li. But also to reduce the original NSL in ^{6}He.

At the same Coulomb repulsion,^{6}Li increases the NSL until it reaches ^{9}Li. Again, β decay diminished the NSL transmuting to ^{9}Be. This element starts again to increase NSL up to ^{10}Be, which again β decayed to ^{10}B to diminish NSL and so on. Hence, every time the surface area per nucleon increases to the unstable limit, β decay occurs to resolve the instability. This produces continuous step decreases all through stable nuclei. The process just described pass through different elements. For example, in Fig. 3 there is an increase in the NSL in the series ^{16}O:^{17}O:^{18}O. Then, there is a small NSL decrease through continuous elements, creating the row ^{18}O:^{19}F:^{20}Ne. This is occurring even though the Coulomb repulsion is increasing. The NSL increases in Ne again, following the series ^{20}Ne:^{21}Ne:^{22}Ne.

Then, another small NSL decrease occurs through elements, forming the row ^{22}Ne:^{23}Na:^{24}Mg with progressive increments in Coulomb repulsion. This is followed by another increase in the NSL in the series ^{24}Mg:^{25}Mg:^{26}Mg. In Fig. 4, the first small decrease in NSL is observed in the row ^{38}Ar: ^{39}K:^{40}Ca. If we follow this row, the next element would be ^{41}Sc. This isotope is unstable because it has too much Coulomb repulsion for the small NSL decrease trade-off. As a consequence, the next stable nucleus occurs in an increase of the NSL, producing ^{40}Ar, which also is accompanied by a significant decrease in Coulomb repulsion. From ^{40}Ar a new row of small decrease of the NSL but progressive increase in Coulomb repulsion starts again, ^{40}Ar:^{41}K:^{42}Ca. This will end at ^{43}Sc, which is unstable for the same reasons discussed above.

Once ^{42}Ca is reached, a new trend of increasing NSL started, ^{42}Ca:^{43}Ca:^{44}Ca. This makes a hole in stability for ^{41}Ca. This isotope is not stable because ^{41}K presented a more favorable trade-off between the NSL and Coulomb repulsion. The next row would be ^{44}Ca:^{45}Sc:^{46}Ti. And the next series ^{46}Ti: ^{47}Ti:^{48}Ti and so on.

^{46}Ca however, appeared as an outlier in this trend. It could be argue that it makes a row with ^{46}Sc but it does not decay to it. It looks like it is an island of NSL stability.

The evidence presented calls to build a model where all nucleons share the surface of the nucleus.

5 Conclusions

The nuclear binding energy is directly related to the nucleons' surface area lost (NSL). A trade-off between the NSL and the Coulomb repulsion is related to the nucleus stability. The progressive increase of the mass in an element will produce different isotopes until its NSL reaches an upper limit for its Coulomb repulsion. Then, β^- decay or neutron emission occur to diminish the NSL and resolve the instability. If there is not enough neutrons (electric insulation) for a given Coulomb repulsion, β^+ decay, proton or α emission occur to diminish it.

References

1. Pomorski K. and Dudek J. P. Nuclear Liquid Drop Model with the Surface-Curvature Terms: New Perspectives for the Hyperdeformation Studies. arXiv: nucl-th/0205011.

2. Möller P. The limits of the nuclear chart set by fission and alpha decay. *EPJ Web of Conferences*, 2016, v. 131, 03002.

3. Angeli I. and Marinova K. P. Table of experimental nuclear ground state charge radii: An update. *Atomic Data and Nuclear Data Tables*, 2013, v. 99, 69.

4. Abrahamyan S. et al. (PREX Collaboration). Measurement of the Neutron Radius of ^{208}Pb through Parity Violation in Electron Scattering. *Phys. Rev. Lett.*, 2012, v. 108, 112502.

5. Pourshahian S. *J. Am. Soc. for Mass Spec.*, 2017, v. 28, 1836.

6. Wang M., Audi G., Kondev F. G., Huang W. J., Naimi S. and Xu X. The AME2016 atomic mass evaluation (II). Tables, graphs, and references. *Chinese Physics C.*, 2017, v. 41 3, 030003.

7. Audi G., Kondev F. G., Wang M., Huang W. J. and Naimi S. The NUBASE2016 evaluation of nuclear properties. *Chinese Physics C.*, 2017, v. 41 3, 030001.

Properties of Superdeformed Rotational Bands in the Perturbed SU(3) Limit of the sdg Interacting Boson Model

A.M. Khalaf[1], Azza O. El-Shal[2], M.M. Taha[2] and M.A. El-Sayed[2]

[1]Physics Department, Faculty of Science, Al-Azhar University Cairo, Egypt. E-mail: ali_khalaf43@hotmail.com
[2]Mathematics and Theoretical Physics Department, NRC, Atomic Energy Authority Cairo, P. No. 13759, Egypt.
Corresponding Author: E-mail: mathelgohary@yahoo.com (M.A. El-Sayed)

The nuclear superdeformed bands in A∼ 190, A∼ 130 mass regions have been systematically analyzed by using the perturbed SU(3) limit of the interacting boson model. The g-bosons have been taken into consideration and the SU(3) symmetry is perturbed by introducing an interaction holding the SO(5) symmetry. A four parameters simple analytic formula for the eigenvalue equation has been derived. The spin determines of the studied superdeformed (SD) bands are considered from our previous works. The improved model parameters for each nucleus have been determined by operating a computer simulated search program so as to obtain a minimum root mean square divergence of the evaluating gamma ray transition energies and the observed ones. With these adopted model parameters the transition energies E_γ, the rotational frequencies $\hbar\omega$, the kinematic $J^{(1)}$ and dynamic $J^{(2)}$ moments of inertia have calculated and are in accordance with experimental data. The behavior of $J^{(1)}$ and $J^{(2)}$ as a function of $\hbar\omega$ have been studied. The calculated E_γ have been used to investigate the anomalous $\Delta I = 2$ staggering by considering the five point formula of Cederwall staggering parameter which represent the finite deviation calculation to the fourth order derivative of the transition energies at a determined spin.

1 Introduction

It was known that the interacting boson model (IBM) [1] with s and d bosons (sdIBM) is successful in studying the spectroscopic properties of low-lying collective states in heavy and medium nuclei. This simple sdIBM allows the utilization of the algebraic symmetries for approaching different type of nuclear spectra, known as dunamical symmetries U(5), SU(3) and O(6) which geometrically describe vibrational, axially deformed and gamma soft nuclei respectively. These three symmetry limits form a Casten triangle [2], that represent the nuclear phase diagram [3]. Transitions of shape phase between these vertices of Casten triangle were widely calculated along several isotopic chains [4–10]. Extended version of IBM where one includes the g-bosons in addition to s and d bosons to account for hexadecapole deformation of the nucleus is receiving a considerable attension of several research groups [11,12]. This hexadecapole deformation is the second most important mulitipolarity in the description of nuclear properties in addition to the quadrupole deformation. An interest in this multipolarity is increased by the observation of the $\Delta I = 2$ energy staggering of superdeformed rotational bands (SDRB's) in some nuclei [13,14], where nuclear spins with rotational sequences splitting by two may divide into two branches. Several theoretical attempts were made for the possible explanation of this $\Delta I = 2$ staggering phenomenon [15–25]. To describe the dynamical symmetries of nuclear states consisting of spdf bosons, it was found [26,27] that one must begin with a supersymmetric group chain U(15,10)

and ending at O(3) due to conservation of angular momentum passing through SU(3) limit of the sdg IBM which is a reasonable starting point to describe SD states in IBM [28]. The sdg IBM is well adopted for study of starting deformed and SD nuclei [15, 16, 26] there is seven different limits of SU(15) [29]. These limits can be splitted into two sets, the first set consists of the three limits which include only partial mixing between the bosons, however the second set consists of four limits which include a mixing of all bosons. If we consider the case of two s, d or g bosons, then the possible angular momenta are $L = 0^3, 2^4, 3, 4^4, 5, 6^2, 8$ where the exponent indicates the multiplicity. The $L = 3, 5$ states are pure dg configurations while the $L = 8$ states is pure g^2. All other states however are mixtures of s, d and g bosons. The difficulty with performing sdg IBM computations for normal deformed and superdeformed nuclei that have boson numbers $N = 12 - 16$ is that the core is too large, and the numerical methods (diagonalization) of the Hamiltonian is not possible. It was proved that the mathematical properties of the $SU(5)_{sdg}$ can be describe the deformed nuclei [30] because by using the intrinsic coherent states [11] the potential energy surface (PES) of the $SU(5)_{sdg}$ limit displays two minima. Since SDRB's are known in the second minimum of the potential well, this property was used [31] to justify an applications of $SU(5)_{sdg}$ limit in SD states. The group SU(3) which relates to the representations $[f_1, f_2, f_3]$ through $\lambda = f_1 - f_2$ and $\mu = f_1 - f_3$ is very important in studying the axial symmetric SDRB's. The one boson state belongs to the $(\lambda, \mu) = (4, 0)$ representation while the two bosons states belongs to $(7,0), (4,2), (0,4)$ rep-

resentation. To appear the $\Delta I = 2$ staggering, the SU(3) must be broken down by adding the $SO(5)_{sdg}$ symmetry as a perturbation. The aim of this work is to use this perturbed SU(3) of sdgIBM to investigate the main properties of superdeformed rotational bands in different nuclei and especially exhibit the $\Delta I = 2$ staggering in their transition energies.

2 Outline of the model

The states of SD bands can be classified in framework of supersymmetric group chain as:

$$U(m,n) \supset U_B(m) \otimes U_F(n) \supset \dots \supset SO_{B+F}(3) \otimes SU_F(\grave{n}) \supset O(3)$$
$$\downarrow \qquad \downarrow \qquad \downarrow \qquad \qquad \downarrow \qquad \downarrow \qquad \downarrow$$
$$[N] \qquad [N_B]_m \qquad [N_F]_n \qquad \qquad L \qquad S \qquad I$$

The notation under those of groups are the corresponding irreducible (irrep) representation. The particles total number $N = N_F + N_B$ with N_F and N_B the fermion and boson numbers respectively. L is the effective core angular momentum and S is the total pseuduspin and I is the total spin of the nucleus. m is determined by the constituent of bosons, while n is determined by the single particle configuration of the fermions and $\grave{n}$ is the total pseudospin. Since the bosons to describe positive parity SD states should be s, d, g bosons [17, 20, 22] and p,f bosons are essential to show negative parity states [27], the space spanned by the single boson states is $\sum_\ell (2\ell + 1) = 1 + 3 + 5 + 7 + 9 = 25$ dimensions. So that, we have the group chain for the boson part

$$U_{sdgpf}(25) \supset U_{sdg}(15) \otimes U_{pf}(10) \supset SU_{sdg}(3) \otimes SU_{pf}(3) \supset SU(3) \supset O(3)$$
$$\downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow \qquad \downarrow$$
$$[N_B] \qquad [N_{sdg}] \qquad [N_{pf}] \qquad (\lambda,\mu)_{sdg} \qquad (\lambda,\mu)_{pf} \qquad (\lambda,\mu) \qquad I$$

The law-lying positive parity states are from the N_{sdg} bosons only, while negative parity states are one pf boson coupled states with $N_{sdg} = N - 1$ sdg bosons. There are also negative parity states formed by coupling odd number of pf bosons with residual sdg bosons and states of positive parity formed by even number of pf bosons with the sdg bosons. Here $N_B = N_{sdg} + N_{pf}$ with $N_{sdg} = 0, 1, 2, \dots, N$ physically N is the number of positive parity bosons. All the irres can be determined with the branching rules [14] of the irres reduction. The reduction $SU(3)_{sdg} \otimes SU(3)_{pf} \supset SU(3)$ can be done in standard Young diagram method [10] and the reduction $SU(3) \supset O(3)$ is the Elliott rule [11]. We notice that for the positive parity states the results of the sdgIBM are still valid. The interaction Hamiltonian of the nucleus corresponding to the above chain takes the form

$$H = \epsilon C_1[U(15)] + kC_2[SU(3)] + cC_2[O(3)] \qquad (1)$$

in which $C_k[G]$ is the k-order Casimir operator of the group G. The energy of the states can be formulated as

$$E(I) = E_0 + \epsilon N + k[\lambda^2 + \mu^2 + \lambda\mu + 3\lambda + 3\mu] + CI(I + 1) \qquad (2)$$

the $C_2[O(3)]$ operator gives the rotational structure. In varaible moment of inertia model [32], the moment of inertia is spin dependent, such that as I increases, the moment of inertia increase due to the antipairing effect. Therefore, Hamiltonian equation (1) can be written as

$$
\begin{aligned}
H = \ & \epsilon C_1[U(15)] + kC_2[SU(3)] \\
& + C_0 \frac{C_2[O(3)]}{1 + f_1 C_2[O(3)] + f_2(C_2[O(3)])^2}
\end{aligned} \qquad (3)
$$

where the terms with f_1 and f_2 take into account many-body interactions which induce antipairing driving and pairing damping effects on the moment of inertia. The energy of the state I in a band considering only the relative excitation of the states in a rotational band is given by

$$E(I) = C_0 \frac{I(I + 1)}{1 + f_1[(I + 1)I] + f_2[(I + 1)I]^2} \qquad (4)$$

To describe the superdeformed rotational bands, we break SU(3) symmetry by adding the symmetry $SO_{sdg}(5)$ as a perturbation to the Hamiltonian. Therefore, the excited energy of the state of positive parity with spin I in SD band is thus given by

$$
\begin{aligned}
E(I) = \ & B[\tau_1(\tau_1 + 3) + \tau_2(\tau_2 + 1)] \\
& + \frac{C_0}{1 + f_1[(I + 1)I] + f_2[(I + 1)I]^2} I(I + 1)
\end{aligned} \qquad (5)
$$

The (τ_1, τ_2) is the irrep of SO(5) group. In practical τ_1, τ_2 being fixed with the branching rules of the irrep reduction as [21–24]

$$(\tau_1, \tau_2) = (\tfrac{I}{2}, 0) \qquad \text{if} \quad I = 4k, 4k + 1 \qquad (k = 0, 1, 2, \dots)$$

$$(\tau_1, \tau_2) = (\tfrac{I}{2} - 1, 2) \quad \text{if} \quad I = 4k + 2, 4k + 3 \quad (k = 0, 1, 2, \dots)$$

3 Analysis of $\Delta I = 2$ staggering in transition energies in SD bands

In framework of collective model [33], the rotational frequency $\hbar\omega$, the kinematic moment of inertia ($J^{(1)}$) and the dynamic moment of inertia ($J^{(2)}$) calculated from γ-ray transition energies for SDRB's are given from the following definitions

$$\hbar\omega = \frac{1}{4}\left[E_\gamma(I + 2 \to I) + E_\gamma(I \to I - 2)\right] \qquad (MeV) \quad (6)$$

$$J^{(2)} = \frac{4}{E_\gamma(I + 2 \to I) - E_\gamma(I \to I - 2)} \qquad (\hbar^2 MeV^{-1}) \quad (7)$$

$$J^{(1)} = \frac{2I - 1}{E_\gamma(I \to I - 2)} \qquad (\hbar^2 MeV^{(-1)}) \quad (8)$$

Table 1: The adopted best model parameters C_0, B, f_1, f_2 obtained from the fitting procedure for the studied SD bands. The bandhead spin I_0 and the experimental lowest transition energy $E_\gamma(I_0 + 2 \rightarrow I_0)$ for each SD is also given.

SD band	I_0 ($\hbar$)	C_0 $\hbar^{-2}$ keV	B keV	f_1 $\hbar^{-2}$	f_2 $\hbar^{-4}$	E_γ (keV)
^{194}Tl(SD1)	14	0.503298E+01	0.18912E-02	0.326365E-03	-0.34134E-03	268.00
^{194}Tl(SD3)	12	0.522016E0+1	0.37473E-01	0.401374E-04	-0.39907E-08	240.50
^{194}Tl(SD5)	10	0.492810E+01	0.36833E-01	0.307779E-04	-0.42746E-08	187.90
^{130}Ce(SD2)	24	0.909181E+01	-0.34824E-02	0.171564E-04	-0.50224E-08	841.00
^{132}Ce(SD1)	30	0.647195E+01	-0.13947E-01	-0.299066E-04	0.34647E-10	808.55
^{132}Nd(SD1)	40	0.419310E+01	0.16107E-01	-0.547523E-04	-0.11468E-10	797.00
^{136}Sm(SD1)	30	0.640396E+01	0.51834E-03	-0.111011E-03	0.17709E-07	888.00
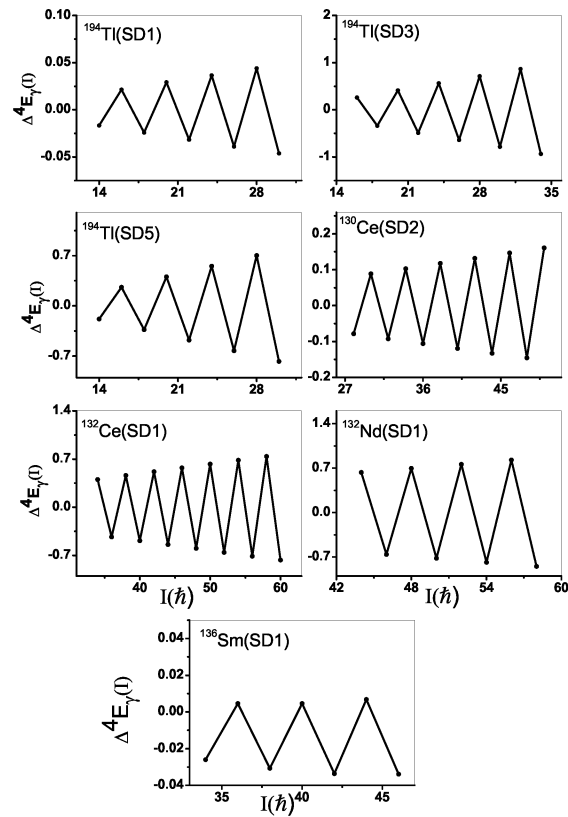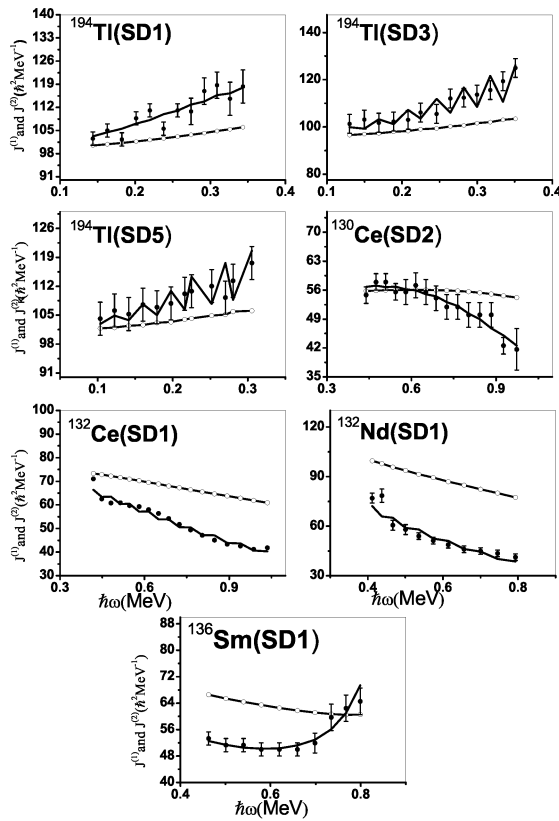
Fig. 1: The calculated results of the kinematic $J^{(1)}$ (open circles) and dynamic $J^{(2)}$ (solid curves) moments of inertia plotted as a function of rotational frequency $\hbar\omega$ for the studied SD bands and the comparison with experimental data for $J^{(2)}$ (closed circles with error bars)

Fig. 2: The calculated $\Delta I = 2$ staggering quantity $\Delta^4 E_\gamma$ obtained by the five point formula as a function of spin for the studied SD bands.

The anomalous $\Delta I = 2$ staggering phenomenon was found in several SD bands [17,18]. Sequences of states which are differing by four units of angular momentum displace relative to each other was shown in superdeformed rotational bands. That is, the SD band can be seen as two sequences of cases with values of spin $I + 4n$ and $I + 4n + 2 (n = 1, 2, 3, ...)$, respectively. This is commonly called $\Delta I = 4$ bifurcation, because the bands divide into two branches with levels differing in spin by $4\hbar$. To explore this $\Delta I = 2$ staggering, the deviation of the $\gamma-$ray energies from a smooth reference $\Delta^4 E_\gamma(I)$ was determined by Cederwall [12], by calculating the finite difference approximation of the fourth order derivation of the $\gamma-$ray energies E_γ at a given spin I by

$$\Delta^4 E_\gamma^{ref}(I) = \frac{1}{16}\Big[E_\gamma(I-4) - 4E_\gamma(I-2) \\ + 6E_\gamma(I) - 4E_\gamma(I+2) + E_\gamma(I+4) \Big] \quad (9)$$

with $E_\gamma(I) = E_\gamma(I) - E_\gamma(I-2)$. The formula (9) contains five energies of consecutive transition and is denoted by the five point formula.

4   Numerical calculations and discussion

For each band of our studied SDRB's, the spin of the bandhead I_0 is taken from our previous works [19–25]. The model parameters C_0, B, f_1, f_2 are determined by using a computer simulated search program in order to obtain a minimum root-mean square (rms) deviation of the calculated transition energies $E_\gamma^{cal}(I)$ from the experimental one $E_\gamma^{exp}(I)$, we employed the common definition of χ

$$\chi = \frac{1}{N} \sqrt{\sum_{i=1}^{N} \left| \frac{E_\gamma^{exp}(I_i) - E_\gamma^{cal}(I_i)}{\delta E_\gamma^{exp}(I_i)} \right|^2} \quad (10)$$

where N is the number of the data points entering into the fitting procedure and $\delta E_\gamma^{exp}(I_i)$ are the experimental errors in $\gamma-$ray energies. Table(1) shows the predicted bandhead spins and the best values of the model parameters C_0, B, f_1, f_2 for each band. Also indicated in Table(1) are the lowest $\gamma-$ray transition energies $E_\gamma(I+2 \rightarrow I_0)$. Using the adopted model parameters, the transition energies E_γ, rotational frequencies $\hbar\omega$, the kinematic $J^{(1)}$ and dynamic $J^{(2)}$ moments of inertia of our selected SD bands are obtained. A very good agreement between the calculated and the experimental values is obtained which gives good support to the model. The kinematic $J^{(1)}$ and dynamic $J^{(2)}$ moments of inertia are plotted as a function of rotational frequency $\hbar\omega$ in Figure(1) compared to the experimental ones. In A~190 mass region, $J^{(1)}$ values are found to be smaller than $J^{(2)}$ and $J^{(2)}$ exhibits a gradual increases with increasing $\hbar\omega$, while in A~130 the values of $J^{(2)}$ are smaller than that the corresponding values of $J^{(1)}$ for all ranges of frequencies and $J^{(2)}$ mostly decrease with a great deal of variation from nucleus to nucleus. Another result in the present work is the observation of a $\Delta I = 2$ staggering effect in $\gamma-$ray energies $E_\gamma(I + 2 \rightarrow I)$ in the studied

SDRB's. The the staggering pattern is illustrated in Figure(2) where the staggering parameters $\Delta^4 E_\gamma(I)$ introduced by Cederwall et al [14] defined as the fourth derivative of E_γ are presented as a function of rotational frequency $\hbar\omega$. A significant zigzag has been observed the resulting numerical values for each band are listed in Tables(2 and 3).

5   Conclusion

The SDRB's namely ^{194}Tl(SD1, SD3, SD5), ^{130}Ce(SD1), ^{132}Nd and ^{136}Sm(SD1) are studied in the version of the perturbed SU(3) limit of sdgIBM with supersymmetry scheme including many body interaction. The bandhead spins are taken from our previous works while the model parameters are adjusted by fitting procedure in order to minimize the relative root mean square deviation between experimental transition energies E_γ^{exp} and the calculated ones E_γ^{cal}. Excellent agreement are given which gives good support to the proposed model. Rotational frequencies, kinematic $J^{(1)}$ and dynamic $J^{(2)}$ moments of inertia are calculated and the evolution of $J^{(1)}$ and $J^{(2)}$ with $\hbar\omega$ are studied. The calculated E_γ are used to investigate the occurrence of a $\Delta I = 2$ staggering effect in the studied SDRB's by using the fourth order derivative of the $\gamma-$ray transition energies. A large amplitude staggering pattern is found in all the studied SDRB's.

References

1. Iachello F., Arima A. The Interacting Boson Model. Cambridge University Press, Cambridge (1987).

2. Casten R.F. Nuclear Structure From a Simple Perespective. Oxford University, Oxford (1990).

3. Warner D. A Triple Point in Nuclei. *Nature*, 2002, v. 420, 614.

4. Arias J.M. et al. U(5)-O(6) Transition in the Interacting Boson Model and the E(5) Critical Point Symmetry. *Phys. Rev.*, 2003, v. C68, 041302.

5. Kusnezov D. The U(16) Algebraic Lattice. *J. Phys.*, 1989, v. A22, 4271, and 1990, v. A23, 5673.

6. Sun H.Z., Zhang M. and Han Q.Z. spdf Interacting Boson Model. *Chin. J. Nucl. Phys.*, 1991, v. 13, 121.

7. Lin Y.X., Sun H.Z. and Zhao E.G. Dynamical Symmetries of the spdf Interacting Boson model. *J. Phys. G*, 1994, v. 20, 407.

8. Khalaf A.M. and Abdelfattah Sh. Structural Evolution in Radium Nuclei Using IBM Consistent-Q- Hamiltonian with Coherent State. *Inernational Journal of Advanced Research in Physical Science*, 2015, v. 2, 24–33.

9. Khalaf A.M. et al., Evolution of Ground State Nuclear Shape in Tungsten Nuclei in Terms of Interacting Boson Model, *ProgressinPhysics*, 2016, v. 13(2), 163–168.

10. Long G.L. The Rotational Limit of the Interacting Boson Model3, *Common Theor. Phys.*, 1997, v. 27, 317.

11. Elliot J.P. Collective Motion of the Nuclear Shell Model I, II. *Proc. Roy. Soc.*, 1958, v. A245 (128), 562.

12. Li S.C. and Kuyucak S. Description of Deformed Nuclei in the sdg Boson Model. *Nucl. Phys.*, 1996, v. A604, 305.

13. Flibotti S. et al. $\Delta I = 4$ Bifurcation in a Superdeformed Band : Evidence for a C_4 Symmetry. *Phys. Rev. Lett.* , 1993, v. 71, 4299.

14. Kota V.K. et al. Group Theoretical Aspects of the Extended Interacting Boson Model. *J. Math. Phys.*, 1987, v. 28, 1644.

Table 2: Calculated γ−ray transition energy E_γ(keV) for the studied SD bands and the comparison with experiment and the calculated results of the $\Delta I = 2$ energy staggering parameter $\Delta^4 E_\gamma(I)$ obtained by five point formula.

^{194}Tl(SD1)				^{194}Tl(SD3)			
$I(\hbar)$	$E_\gamma(exp)$	$E_\gamma(cal)$	$\Delta^4 E_\gamma(I)$	$I(\hbar)$	$E_\gamma(exp)$	$E_\gamma(cal)$	$\Delta^4 E_\gamma(I)$
14	268.0	268.6751		12	240.5	238.2344	
16	307.0	307.4222		14	280.0	278.2792	
18	345.1	345.7051	-0.0165531	16	318.8	318.5983	0.262866
20	384.2	383.6179	0.0213944	18	358.1	357.3467	-0.33660
22	421.0	420.9822	-0.0241556	20	39.2	396.8819	0.412796
24	457.0	457.9697	0.0290825	22	425.3	434.1794	-0.486442
26	494.9	494.3612	-0.0316294	24	473.0	472.8180	0.562820
28	530.9	530.4057	0.0365431	26	510.9	508.5947	-0.636306
30	567.0	565.8415	-0.0389094	28	546.6	546.3097	0.712723
32	601.2	600.9968	0.0438844	30	582.2	580.5822	-0.786148
34	634.9	635.5733	-0.0463116	32	617.4	617.4370	0.862580
36	669.8	669.9782		34	652.0	650.3192	-0.935853
38	703.6	703.8759		36	685.5	686.4767	
				38	717.5	718.1790	

^{194}Tl(SD5)				^{130}Ce(SD2)			
$I(\hbar)$	$E_\gamma(exp)$	$E_\gamma(cal)$	$\Delta^4 E_\gamma(I)$	$I(\hbar)$	$E_\gamma(exp)$	$E_\gamma(cal)$	$\Delta^4 E_\gamma(I)$
10	187.9	186.6115		24	841.0	842.347	
12	226.3	225.4885		26	914.0	912.752	
14	264.0	263.6266	-0.183757	28	983.0	982.836	-0.0785506
16	302.0	302.1677	0.258395	30	1052	1053.335	0.0890181
18	339.2	339.3166	-0.330951	32	1124	1123.728	-0.0921319
20	376.6	377.4083	0.405704	34	1196	1194.918	0.10318
22	413.7	413.4865	-0.478076	36	1266	1266.335	-0.105713
24	450.0	451.0842	0.553079	38	1338	1339.059	0.117491
26	486.1	486.0854	-0.625427	40	1412	1412.478	-0.119078
28	521.8	523.2236	0.700607	42	1489	1487.861	0.132018
30	558.4	557.2253	-0.772767	44	1566	1564.571	-0.132682
32	593.7	594.0268		46	1646	1644.083	0.146972
34	627.7	627.2002		48	1726	1725.751	-0.145203
				50	1806	1811.279	0.161256
				52	1900	1900.046	
				54	1996	1994.015	

15. Hamamoto I. and Mottelson B. Superdeformed Rotational Bands in the Presence of Y_{44} Deformation. *Phys. Lett.*, 1994, v. B333, 294.

16. Pavlichenkov I.M. and Flibotte S. C_4 Symmetry and Bifurcation in Superdeformed Bands. *Phys. Rev.*, 1995, v. C51, R460.

17. Kota V.K.B. Interacting Boson Model Basis and Hamiltonian for $\Delta L = 4$ Staggering *Phys. Rev. C*, 1996, v. 53, 2550.

18. Joki H. and Wu L.A. $\Delta I = 4$ Bifurcation in Ground Bands of Even-Even Nuclei and the Interacting Boson Model. *Phys. Rev. Lett.*, 1997, v. 79, 2006.

19. Khalaf A.M., Taha M.M. and Kotb M. Identical Bands and $\Delta I = 2$ Staggering in Superdeformed Nuclei in A∼ 150 Mass Region Using Three Parameters Rotational Model. *Progress in Physics*, 2012, v. 8, 39.

20. Otsuka T. and Honma M. Interacting Boson Model For Superdeformation. *Phys. Lett. B*, 1991, v. 268, 305.

21. Khalaf A.M. and Okasha M.D. Properties of Nuclear Superdeformed Rotational Bands in A∼ 190 Mass Region. *Progress in Physics*, 2014, v. 10, 246.

22. Kuyucak S., Honma M. and Otsuka T. Description of Superdeformation Nuclei in the Interacting Boson Model. *Phys. Rev. C*, 1996, v. 53, 2194.

23. Khalaf A.M., Awaad T.M. and Elgabery M.F. Examination of $\Delta I = 2$ Energy Staggering in the Superdeformed Bands of ^{194}Hg. *International Journal in Pysical and Applied Science*, 2016, v. 3(4), 47.

24. Khalaf A.M., Sirage M.M. and Abdelmageed K.E. Properties of Superdeformed Bands in A∼ 190 Region Within the Framework of Three Parameters Nuclear softness Model. *Chinese Journal of Physics*, 2016, v. 54, 329.

Table 3: Same as Table 2 but for ^{132}Ce(SD1), ^{132}Nd(SD1) and ^{136}Sm(SD1).

^{132}Ce(SD1)				^{132}Nd(SD1)			
$I(\hbar)$	$E_\gamma(exp)$	$E_\gamma(cal)$	$\Delta^4 E_\gamma(I)$	$I(\hbar)$	$E_\gamma(exp)$	$E_\gamma(cal)$	$\Delta^4 E_\gamma(I)$
30	808.55	804.9409		40	797.0	793.4394	
32	864.85	865.2221		42	849.0	848.8271	
34	928.80	928.3618	0.404878	44	900.0	909.5820	0.630371
36	994.63	991.3491	-0.431978	46	966.0	971.1123	-0.658327
38	1060.32	1057.651	0.460603	48	1035.0	1038.912	0.69562
40	1127.27	1123.824	-0.487641	50	1109.0	1107.942	-0.721866
42	1194.72	1193.792	0.516664	52	1187.0	1184.293	0.760989
44	1263.63	1263.680	-0.543336	54	1269.0	1262.507	-0.785341
46	1334.56	1337.876	0.572326	56	1356.0	1349.300	0.826478
48	1408.34	1412.078	-0.599181	58	1445.0	1438.824	-0.847901
50	1452.67	1491.137	0.629027	60	1537.0	1538.453	
52	1566.70	1570.322	-0.655639	62	1634.0	1641.996	
54	1651.49	1654.962	0.685363				
56	1740.29	1739.900	-0.71106				
58	1832.64	1830.941	0.740601				
60	1926.50	1922.515	-0.766111				
62	2023.50	2020.902					
64	2119.00	2120.123					

^{136}Sm(SD1)			
$I(\hbar)$	$E_\gamma(exp)$	$E_\gamma(cal)$	$\Delta^4 E_\gamma(I)$
30	888	887.0356	
32	963	963.2182	
34	1041	1041.0786	-0.0260325
36	1119	1120.2441	0.0045813
38	1199	1199.9257	-0.0306712
40	1279	1279.4072	0.00451688
42	1359	1357.4814	-0.0336294
44	1436	1433.0136	0.00677437
46	1503	1504.3310	-0.0338756
48	1567	1569.8691	
50	1629	1627.5214	

25. Khalaf A.M., Okasha M. and Ragheb E. Anomalous $\Delta I = 2$ Transition Energy Staggering in Nuclear Superdeformed Bands. *Australian Journal of Basic and Applied Science*, 2016, v. 10, 192.

26. Liu Y.X. and Gao D.F. Description of Identical Superdeformed Bands with $\Delta I = 4$ Bifurcation. *Phys. Rev.*, 2001, v. C63, 044317.

27. Liu X.Y., Wang J.J. and Han Z. Description of the Identical Superdeformed Bands and $\Delta I = 4$ Bifurcation in the A~ 130 Region. *Phys. Rev. C*, 2001, v. 64, 064320.

28. Wu L.A. and Toki H. Anamalous Staggering in the Mass of the Rotational Nuclei in the Actinide Region. *Phys. Lett. B*, 1997, v. 407, 207; $\Delta I = 4$ Bifurcation in Ground Bands of Even-Even Nuclei and the Interacting Boson Model. *Phys. Rev. Lett.*, 1997, v. 79, 2009.

29. Demeyer H. et al. Classification of the Dynamical Symmetries in the Extended Interacting Boson Model. *J. Phys. A: Math. Gen.*, 1986, v. 19, L565.

30. Sun H.Z. et al. SU(3) and SU(5) Dynamical Symmetries in The Extended Interacting Boson Model. *Kinam Rev. Fis.*, 1983, v. 5, 135.

31. Liu Y.X., Sun H.Z. and Zhao E.G. Description of Superdeformed Nuclear States in the Interacting Boson Model. *Commun. Theor. Phys.*, 1997, v. 27, 71.

32. Bonatsos D. and Kelin A. Generalized Phenomenological Models of the Yrast Band. *Phys. Rev. C*, 1984, v. 29, 1879.

33. Bohr A. and Mottelson B. Nuclear Structure. Benjamin, New York, 1975, Vol. II.

A Solution to the Flyby Anomaly Riddle

Eduardo D. Greaves[1], Carlos Bracho[2], and Imre Mikoss[3]

[1]Universidad Simón Bolívar. Apartado 89000, Caracas, Venezuela. E-mail: egreaves20002000@yahoo.com
[2]Facultad de Ingeniería, Universidad Central de Venezuela, Caracas, Venezuela. E-mail: bracho_carlos@hotmail.com
[3]Universidad Simón Bolívar. Apartado 89000, Caracas, Venezuela. E-mail: imikem@gmail.com

The Flyby Anomaly is one of the unsolved problems of current physics in that the Doppler-shift determined speeds are inconsistent with expected values assuming the validity of Newtonian gravity. We postulate that the Flyby Anomaly is a consequence of the assumption that the speed of light is isotropic in all frames, and invariant in the method used to measure the velocity of the space probes by means of the Doppler Effect. The inconsistent anomalous values measured: positive, null or negative are simply explained relaxing this assumption. During space probe energy assistance maneuvers the velocity components of the probe in the direction of the observer V_o are derived from the relative displacement Δf of the radiofrequency f transmitted by the probe, multiplied by the local speed of the light c' by the Doppler effect: $V_o = (\Delta f/f)\,c'$. According to the Céspedes-Curé hypothesis, the movement through variable gravitational energy density fields produces slight variations of the refractive index n' of space and therefore of the speed of light c' which leads to unaccounted corrections of the Doppler data that are based on an invariant c. This leads to incorrect estimates of the speed or energy change in the flyby maneuver in the Earth's frame of reference. The simple theory presented is applied to hyperbolic flyby trajectories of Galileo I and the spacecraft NEAR accurately reproducing the NASA measured values and thereby providing additional experimental evidence for a variable speed of light dependence on the gravitational energy density of space with fundamental consequences in astrophysics and cosmology.

1 Introduction

The Flyby Anomaly is an unexpected energy increase or decrease of spacecraft during flybys maneuvers of Earth and other planets employed as gravitational assist techniques for Solar system exploration. The anomalous measurements have been observed as shifts in the S-band and X-band Doppler and ranging telemetry. It has been observed in a number of spacecraft: NEAR, Galileo I and II, Cassini, Rosetta I, II and III, Messenger, Juno, Hayabusa, and EPOXI I and II [1–3]. The Flyby Anomaly has been included in a list of "unsolved problems in physics". We find very significant a comment of Anderson et al. [2], that the same inconsistency in the Doppler residuals which lead to the velocity anomaly are found in the ranging data, as we believe both can be explained by the theory developed here.

A large number of papers have been advanced in attempts to explain the anomalous, and at times inconsistent, measurement results of the very small, but significant, unaccounted speed and energy change experienced by spacecraft during maneuvers to increase or decrease its relative energy.

A comprehensive review of anomalous phenomena observed in the solar system was published by Lämmerzahl et al. (2006) [4] which includes prominently the Flyby Anomaly. It lists numerous possible causes of the anomaly. It reaches the conclusion, in this respect, that none of them can explain the observed measurements. "New physics" has been attempted by postulating variants of gravitational theories [5–9], or modification of inertia [10], and also the possible influence of halos of dark matter [11].

More conventional causes that have been considered include: The effect of Earth oblateness which is known to produce perturbations of orbiting spacecraft. Hence a possible cause of the Flyby Anomaly might be the non spherical mass distribution of the oblate Earth. An unsuccessful attempt has been made by K. Wilhelm and B.N. Dwivedi (2015) [12] to explain the anomalous Earth flybys of several spacecraft on the basis of asymmetry of the mass distribution of the Earth causing an offset of the effective gravitational centre from the geometric centre.

The possibility of electromagnetic forces acting between a charged probe and the Earth's magnetic fields has been examined [13], also the influence of the Earth high atmosphere [14] or the emission of thermal energy from the spacecraft [15]. However, to this date none of the above adequately explains the cause of the anomaly.

A light speed anisotropy hypothesis is used by R.T. Cahill to argue that the Doppler-shift determined speeds are inconsistent with expected speeds, and hence affect the measurement of the probe during flyby [16]. Cahill revisits the Michelson-Morley experiment controversy citing numerous new interferometer results which take into account the effect if the medium that light transverses in these experiments (e. g. gas, coaxial cable or optical fiber). He points out that speed anomalies are not real and are actually the result of using an incorrect isotropic light speed relationship between the observed Doppler shift and the speed of the spacecraft.

An empirical formula that adequately predicts the flybys measured up to 2005 was published by Anderson et al. [1, 2] using all likely variables in the problem. The empirical formula developed by Anderson et al. did not fit later anomalous flybys. However, a modification by Jouannic et al. (2015) [3] was able to predict the new data. From the conclusions of this work we read that "This could signify that it (*the anomaly*) is caused by a force related either to mass, altitude, or both". In this paper we show that indeed, planet mass and distance from the planet, which are some of the important variables in determining the gravitational energy density of space and hence of the local index of refraction of quasi-empty space [17, 18] produces minute variations in the local speed of light c' due to the Céspedes-Curé hypothesis [19], explained below. These unaccounted variations of the local index of refraction lead to small erroneous measurements of spacecraft velocity and derived energy, based on a constant c, and is shown here to be the cause of the Flyby Anomaly. Hence we coincide with Cahill in that speed anomalies are not real but rather an artifact of how the speeds are measured with the Doppler effect. In this paper the fundamentals of the proposed Flyby Anomaly explanation are presented with analytical relations showing how the anomalous behavior can be accurately predicted. Numerical calculations are presented for the Galileo I (December, 1990) Earth flyby and NEAR (January, 1998) Earth flyby. We also show how the anomaly can be simply predicted for any other spacecraft provided detailed information of the measurement of entry and exit points are available. Additionally we briefly discuss some of the fundamental consequences of the Céspedes-Curé hypothesis for astrophysics and cosmology.

2 Speed and energy measurement of spacecraft and the Doppler effect

All remote velocity estimations of astronomical bodies use the first order Doppler effect of light [20]. In spacecraft the procedure employs a locally produced radio or light frequency f of accurately known value, or it could be a retransmitted signal such as the case of Pioneer spacecraft [21]. The speed component in the direction of the observer V_o is deduced from the shift Δf of the radio or light frequency f, times the local speed of light c' by means of $V_o = (\Delta f / f) c'$. At the present time (year 2020) it is conventionally assumed that the local speed of light c' at any point in the universe is isotropic and identical to the speed of light $c = 299792458 \, \text{ms}^{-1}$ measured in vacuum to high accuracy on the surface of the Earth. Clearly, if there are small variations of c' as a result of changing locations with differing gravitational energy density ρ, as occurs during flyby maneuvers, the measured speed component in the direction of the observer V_o, calculated with the Doppler effect, assuming a constant c, will lead to erroneous estimations of the spacecraft speed and resulting energy change during the maneuver. Presently the speed of light

c is considered a fundamental constant being the base of the definition of the meter, the length unit in the SI system of units. However, a variable speed of light has been considered by a number of authors, notably including A. Einstein in 1907 [22] and in 1911 [23] and also by R. Dicke in 1957 [24]. In Einstein's early work the speed of light was influenced by the gravitational potential and a constant speed could not be conceived in a gravitational field with variable strength. In Dicke's work he assumes a refractive index n of empty space, different from 1, given by an expression where the value increases with the gravitational field:

$$n = 1 + \frac{GM}{r c^2}.$$

This proposal provides an alternative to the lensing phenomenon predicted by General Relativity Theory (GRT). There are other more modern variable speed of light theories as reviewed by Magueijo J. in 2003 [25]. The Céspedes-Curé hypothesis [19] is reminiscent of the early proposals of Einstein and Dicke. It predicts that the speed of light is a function of the local total energy density of space ρ according to (1), so that if this hypothesis is correct, it could explain the spacecraft anomalous behavior derived by the Doppler effect.

$$c = \frac{k}{\sqrt{\rho}}, \tag{1}$$

where k is a proportionality constant and ρ is the sum of all the sources of energy density including gravitational, ρ_G, electric, ρ_E, magnetic, ρ_M, and any other that may be acting at the site. Calculations [26] show that gravitational energy density is much larger than electric or magnetic. And that the most important source of energy density by several orders of magnitude is the "Cosmic energy density" due to the far away stars and galaxies which has a value of $\rho^* = 1.094291 \times 10^{15} \, \text{Jm}^{-3}$ deduced by Céspedes-Curé [19], see Appendix A, and by Greaves E.D. [18, 26, 27], see Appendix B. Compared to ρ^*, the Sun's ρ_S, the planet about which the flyby maneuver is being done, ρ_p, and all other massive bodies in the vicinity contribute in a very minor amount to the variable total energy density at points along the trajectory of the spacecraft. Hence, this is the cause of the minute amount found for the anomalous values of velocity and energy of spacecraft performing the flyby maneuver. The gravitational energy density ρ due to a mass M at a distance r from its center is given by [19, see page 163],

$$\rho = \frac{1}{2} \frac{GM^2}{4\pi r^4} = \frac{GM^2}{8\pi r^4}, \tag{2}$$

where G is the universal constant of gravitation. Using this relation the gravitational energy density of any astronomical mass can be calculated at any point in space located a distance r from the mass center. The energy density of space ρ_B and ρ_E associated with the presence of static magnetic B and electric

E fields are given by [28]:

$$\rho_B = \frac{1}{2\mu_0}\, B^2 \,, \tag{2a}$$

and

$$\rho_E = \frac{1}{2}\varepsilon_0\, E^2 \,, \tag{2b}$$

where μ_0 is the magnetic permeability and ε_0 is the electric permittivity of free space. With the usual definition of the index of refraction at a point in space, n', as the ratio of the speed of light of vacuum c on the surface of Earth to the speed of light c' at the point considered (conventionally inside a transparent material) $n' = c/c'$ it is possible with the use of (1) to obtain a relation for n' which is only dependent on values of the energy density of space at the point in question and at the surface of the Earth:

$$n' = \frac{c}{c'} = \frac{\sqrt{\rho'}}{\sqrt{\rho}} = \frac{\sqrt{\rho'}}{\sqrt{\rho^* + \rho_S + \rho_E}}. \tag{3}$$

Here $\rho^* + \rho_S + \rho_E$ is the gravitational energy density at the surface of the Earth. The terms in the sum are: the energy density due to the far away stars and galaxies ρ^*, the Sun, ρ_S and Earth, ρ_E. The values shown in Table 1 and Fig. 1 indicate that the contributions to the local gravitational energy density due to nearby planets is small and negligible compared to the all-pervading energy density ρ^* due to the far away stars and galaxies. Hence for a spacecraft in a flyby maneuver the local value of the index of refraction n' and the local value of the speed of light c' is very nearly equal to the values on the surface of Earth. This leads to the fact that the observed anomalous variations of the speed of spacecraft deduced by the Doppler effect are very small indeed. It also shows that the anomalies are dependent on the mass of the planet and on the distance to the planet as mentioned in the conclusions of the work of Jouannic et al. in [3].

3　Calculation of the anomaly

In order to predict quantitatively the measured energy change that shows an anomalous value it is necessary to have very detailed information of the particular flyby event considered. The information required is data that refers to the spacecraft such as the radio frequencies used for transmission which are used for determining the relative radial velocity via the Doppler effect. The information related to the planet, about which the maneuver takes place, is information that defines the orbit of the spacecraft: the hyperbolic orbit parameters of the flyby: a (semi-mayor axis) and e (eccentricity) and the entry and exit velocity of the probe: V_∞^- and V_∞^+, the measured anomalous velocity V_{anom} and, most important, the points of entry and exit where the velocities were measured. NASA determines the Flyby Anomaly with the Orbit Determination Program (ODP) of the Jet Propulsion Laboratory
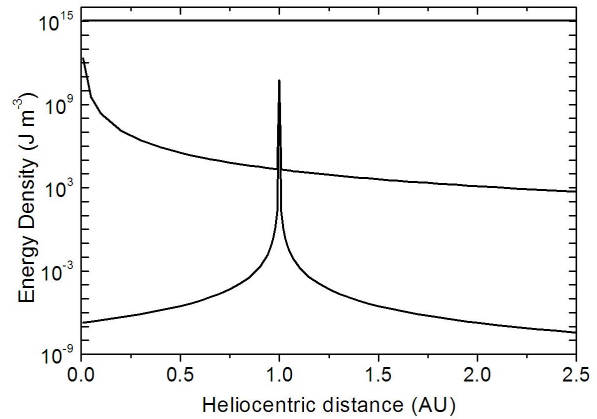

Fig. 1: Gravitational energy density (Jm^{-3}) as a function of distance from the center of the Sun in AU (0 to 2.5 AU) due to the far away stars and galaxies (top line $\rho^* = 1.094291 \times 10^{15}$ Jm^{-3}), ρ_S due to the Sun (middle line) and ρ_E due to Earth (Line centered at 1 AU) [26].

(JPL) as well as other software at the Goddard Space Flight Center and at the University of Texas [2]. These programs incorporate all the physics mentioned above and the information gathered by the Deep Space Network (DSN) during the flyby. According to the hypothesis presented in this paper the anomaly is due to errors committed due to sub-estimation or over-estimation of the velocity calculated by the use of the Doppler effect formula as explained previously. Below we show how the anomaly can be calculated in reference to Earth flybys. The same considerations apply to flybys about other planets. From (3) we derive

$$c' = c\, \frac{\sqrt{\rho}}{\sqrt{\rho'}}\,. \tag{4}$$

The radial velocity of the spacecraft during the flyby is obtained by the use of $V_r = \Delta f/f\; c'$ which with (4) gives

$$V_r = c'\, \frac{\Delta f}{f} = c\, \frac{\Delta f}{f}\, \frac{\sqrt{\rho}}{\sqrt{\rho'}}\,, \tag{5}$$

where the gravitational energy density ρ' is a function of the position of the spacecraft in its orbit and ρ is the gravitational energy density on the surface of the Earth whose value is $\rho = \rho^* + \rho_S + \rho_E$ with ρ_S and ρ_E calculated on the surface of Earth. As the spacecraft nears the planet it moves into varying values of ρ' which according to (5) results in a sub-estimation or over-estimation of the velocity. Likewise, as the spacecraft leaves the vicinity of Earth and gets further away, it travels into different values of the gravitational energy density ρ' which according to (5) results in differing values of the velocity. Important factors determining the value of ρ' are the radial distance to the center of the planet producing the energy

Table 1: Values of the energy density of space at the surface of Earth produced by: the far away stars and galaxies, the mass of the Sun, Earth, the Moon and other planets.

Source of energy density	Symbol	Energy density due to source at	Magnitude (Joules/m^3)‡	Reference
Far away Stars and Galaxies	ρ^*	Earth	1.094291×10^{15}	Céspedes-Curé [19, p. 279]
Sun	ρ_S	Earth 1 AU	2.097×10^4	Greaves [17, 18]
Sun	$\rho_{S@AU-}$	1 AU$-E_{\text{SI}}{}^\dagger$	2.150250×10^4	This work
Sun	$\rho_{S@AU+}$	1 AU$+E_{\text{SI}}{}^\dagger$	2.046034×10^4	This work
Earth	ρ_E	Earth surface	5.726×10^{10}	Greaves [18]
Moon	ρ_{Moon}	Earth	6.57×10^{-1}	Greaves [18]
Jupiter	ρ_{Jup}	Earth	1.91×10^{-2}	Greaves [18]
Venus	ρ_{Ven}	Earth	2.14×10^{-5}	Greaves [18]
Mars	ρ_{Mar}	Earth	2.91×10^{-8}	Greaves [18]

† E_{SI} is the radius of Earth's Gravitational Sphere of Influence: (929000 km) [29, 30].
‡ These values are deceptive due to the $1/r^4$ dependence of the gravitational energy density (2). The energy density of the Earth at its surface is 6 orders of magnitude greater than the Sun's. However, it decreases abruptly so that at a distance greater than 41 earth radii the energy density due to the Sun is higher.

assistance and the radial distance to the Sun. Hence, in order to calculate exactly the anomalous energy change reported, it is necessary to know the exact position of the spacecraft at the point or points where its velocity was calculated in order to establish the initial spacecraft energy and the point or points where the velocity was finally calculated to establish the final spacecraft energy. Also needed are the methods used for the speed measurements such as the frequency used by the spacecraft in its transmission to the Earth tracking stations, and whether it is a spacecraft transmission or an Earth sent-signal retransmitted by the spacecraft. Such detailed information is ordinarily not included in papers publicly available.

Examination of (5) shows that the anomaly is caused by the square root term (SQR)

$$\text{SQR} = \sqrt{\frac{\rho}{\rho'}} = \sqrt{\frac{\rho}{\rho^* + \rho_S + \rho_E}}. \tag{6}$$

Here ρ and ρ^* are constants while ρ_S and ρ_E are functions of position, ρ_S is dependent on the radial distance to the center of the Sun and ρ_E is dependent on the radial distance to the center of Earth.

Let us consider ρ_S first, which is given by

$$\rho_S = \frac{GM_S^2}{8\pi r_S^4}. \tag{7}$$

Here M_S is the mass of the Sun and r_S the radial distance from the center of the Sun. In order to estimate the influence of this term we calculate the value of ρ_S over the Earth's gravitational Sphere of Influence, E_{SI}, that is at a distance of one

AU from the Sun in the range of 1 AU $\pm E_{\text{SI}}$ (plus or minus the radius of the Earth's Sphere of Influence). The values obtained range from $\rho_S = 2.150250 \times 10^4$ to 2.046034×10^4 Jm^{-3} as shown in Table 1. The variation over the Earth's sphere of influence is of the order of 5%. However, the values of the variation of the gravitational energy density due to the Sun are 5 orders of magnitude less than the energy density due to Earth at its surface. But, as shown by calculations, they become more important than the Earth's energy density due to the $1/r^4$ term in (2) as discussed below.

In (6), the value of ρ_E is given by

$$\rho_E = \frac{GM_E^2}{8\pi r_E^4} \tag{8}$$

with M_E the mass of the Earth and r_E the radial distance from the center of Earth.

Taking these considerations into account in (5) we can write an expression for the corrected speed of the spacecraft which takes into account the change of the index of refraction of space due to the variation of the space gravitational energy density along the spacecraft trajectory:

$$\begin{aligned} V_r &= c \frac{\Delta f}{f} \sqrt{\frac{\rho}{\rho^* + \rho_S + \rho_E}} \\ &= c \frac{\Delta f}{f} \sqrt{\frac{\rho}{\rho^* + \frac{GM_S^2}{8\pi r_S^4} + \frac{GM_E^2}{8\pi r_E^4}}}. \end{aligned} \tag{9}$$

Numerical calculations show that the influence of the third

term of the denominator, namely the variation of the Earth's gravitational energy density is important only at small distances above the surface of the Earth and it becomes very small at distances where a spacecraft is beginning its approach to the surface of the planet during a flyby.

4 Calculation of the Flyby Anomaly in three cases

To calculate the anomaly, we suppose that the speed of the spacecraft is measured at two points: a point of entry into the Earth's sphere of influence where the speed is V_∞^- and a point of exit from the Earth's sphere of influence where the speed is V_∞^+. If we ignore the change of c, the measured velocities are given by:

$$V_\infty^+ = c \frac{\Delta f^+}{f} \quad \text{and} \quad V_\infty^- = c \frac{\Delta f^-}{f}.$$

Hence the anomaly measured by NASA is given by

$$\text{An} = V_\infty^+ - V_\infty^- = \frac{c}{f} \left(\Delta f^+ - \Delta f^- \right). \tag{10}$$

At each of these points a correct measurement, one that takes into account the change of the index of refraction, as we propose in this paper, must be done with (9), with V_∞^- the observed Doppler shift at the point of entry, and with V_∞^+ the observed Doppler shift at the point of exit as shown below:

$$V_\infty^+ = c \frac{\Delta f^+}{f} \sqrt{\frac{\rho}{\rho^* + \frac{GM_S^2}{8\pi(r_S^+)^4} + \frac{GM_E^2}{8\pi(r_E^+)^4}}} \tag{11a}$$

$$V_\infty^- = c \frac{\Delta f^-}{f} \sqrt{\frac{\rho}{\rho^* + \frac{GM_S^2}{8\pi(r_S^-)^4} + \frac{GM_E^2}{8\pi(r_E^-)^4}}} \tag{11b}$$

In the Earth's coordinate system, energy is conserved, so that if the correct equations (11a) and (11b) are used, then measurements *should* give: $V_\infty^+ - V_\infty^- = 0$ that is:

$$0 = c \frac{\Delta f^+}{f} \sqrt{\frac{\rho}{\rho^* + \frac{GM_S^2}{8\pi(r_S^+)^4} + \frac{GM_E^2}{8\pi(r_E^+)^4}}}$$
$$- c \frac{\Delta f^-}{f} \sqrt{\frac{\rho}{\rho^* + \frac{GM_S^2}{8\pi(r_S^-)^4} + \frac{GM_E^2}{8\pi(r_E^-)^4}}}. \tag{12}$$

However, if the SQR terms are different, for (12) to be true it requires that $\Delta f^+ \neq \Delta f^-$, and hence measurements done by NASA with (10) will show an anomaly. The anomaly is contained in the difference of the SQR terms in (12). Since

$$V_\infty^+ = c \frac{\Delta f^+}{f} \quad \text{and} \quad V_\infty^- = c \frac{\Delta f^-}{f}$$

are almost the same, both of the order of km/s differing by an amount 6 orders of magnitude smaller, of the order of mm/s,

we can write the following relation to calculate the measured anomaly:

$$V_{\text{anom}} = V_\infty \sqrt{\frac{\rho}{\rho^* + \frac{GM_S^2}{8\pi(r_S^+)^4} + \frac{GM_E^2}{8\pi(r_E^+)^4}}} -$$
$$\tag{13}$$
$$- V_\infty \sqrt{\frac{\rho}{\rho^* + \frac{GM_S^2}{8\pi(r_S^-)^4} + \frac{GM_E^2}{8\pi(r_E^-)^4}}}.$$

Numerical analysis of (13) shows it is possible to identify three cases.

4.1 First case

The distances from the point of entry and the point of exit to the Sun and to Earth are the same. ($r_S^+ = r_S^-$ and $r_E^+ = r_E^-$). In this case the two terms in the parenthesis of (13) are the same and no anomaly will be detected (incoming and outgoing points are symmetric with respect to the Sun and Earth).

4.2 Second case

In this second case entry point and the exit point are at different distances from the Sun but at the same distance from Earth. It means that $r_S^+ \neq r_S^-$, hence:

$$\frac{GM_S^2}{8\pi(r_S^+)^4} \neq \frac{GM_S^2}{8\pi(r_S^-)^4},$$

so that the SQR terms in (12) are different. For this relation to be correct it requires that $\Delta f^+ \neq \Delta f^-$. Hence if the speeds are being measured with relations

$$V_\infty^+ = c \frac{\Delta f^+}{f} \quad \text{and} \quad V_\infty^- = c \frac{\Delta f^-}{f}$$

as in (10) the flyby will certainly show an anomaly: $V_\infty^+ \neq V_\infty^-$. However, numerical calculations show that the anomalous values in this case are very small and non measurable.

4.3 Third case

In this third case entry point and the exit point are at different distances from the Sun and at different distance from Earth. It means that, $r_S^+ \neq r_S^-$ and $r_E^+ \neq r_E^-$. In this case the two terms in the parenthesis of (13) are different. Hence if the speeds are being measured with relations

$$V_\infty^+ = c \frac{\Delta f^+}{f} \quad \text{and} \quad V_\infty^- = c \frac{\Delta f^-}{f}$$

as in (10) the flyby will certainly show an anomaly: $V_\infty^+ \neq V_\infty^-$. Numerical calculations show that an anomaly will be measured in the range of values reported, negative or positive, with a value and sign that depends on the entry and exit points used for measurement. We conclude that the anomaly is due to neglect of the SQR terms in the calculation of the entry and exit velocities derived from the Doppler flyby data.

Table 2: Distances to the Sun and to Earth with calculated entry and exit points that predict, with (13), the measured Flyby Anomaly of the Galileo 1 (December 1990) flyby and the NEAR (January 1998) flyby.

	Galileo 1		NEAR	
	Entry point	Exit point	Entry point	Exit point
Distance from Sun (m)	1.502803×10^{11}	1.502831×10^{11}	1.495630×10^{11}	1.495950×10^{11}
Distance from Earth (m)	1.7651×10^{7}	1.4864×10^{7}	7.2000×10^{7}	1.2200×10^{7}
Spacecraft Velocity (m/s)	8949		6851	
Measured Flyby Anomaly (mm/s)	3.930		13.46	
Calculated Flyby Anomaly (mm/s)	3.944		13.38	
Difference (%)	+0.40		−0.57	

5 Results

In order to apply the theory described above to predict the anomaly measured for any given spacecraft flyby it is necessary to introduce into (13) the values of the parameters of the spacecraft maneuver, namely the spacecraft speed at the entry point and the distances to the Sun and to Earth of the incoming and outgoing points. The spacecraft speed is available, however, the required information of entry and exit points has not been possible to obtain. Only the right ascension and declination of these vector directions are given by Anderson et al. [2]. With these angular parameters we have defined vectors, from the Earth, for incoming and outgoing directions as well as from the Earth to the Sun's direction along its right ascension and declination on the day of the Flyby. Then with calculated tables of numerical values of the SQR terms of (13) for varying entry and exit points along the incoming and outgoing vectors (i.e. values of r_S^+, r_S^- and of r_E^+, r_E^-) excluding the immediate distances (1h 40min before and after the closest approach location) we have arrived at likely entry and exit points that closely predict the observed NEAR (January 23, 1998) flyby. For Galileo I (December 8, 1990) flyby the incoming and outgoing points were calculated along likely in and out points not specifically along the actual incoming and outgoing vectors. Results of these calculations are shown in Table 2.

6 Possible measurement of ρ^* with the Flyby Anomaly

Based on the Flyby Anomaly explanation given above, it is possible to use the experimental results of measured flyby anomalies in spacecraft to calculate, in an independent way, the gravitational energy density values that lead to the measured anomalies. Since the gravitational energy density is composed of the contribution due to the planets and the Sun, which can be accurately calculated with (8), the contribution due to the far away stars and galaxies, ρ^*, could be solved as a single adjustable parameter, and calculated. This could be done by programming the theory presented here in the Orbit Determination Program of the JPL, or by an accurate knowledge of the points of entry and exit in the hyperbolic trajectory

where the measurements were made that produced a Flyby Anomaly. This measurement of ρ^*, the gravitational energy density of the far away stars and galaxies, would provide an additional estimation of its value besides that given by Jorge Céspedes-Curé [19, page 279], $\rho^* = 1.094291 \times 10^{15}$ Jm^{-3}, obtained using starlight deflection measurements during total sun eclipses, see Appendix A, or that given by Greaves [26]: $\rho^* = 1.0838 \times 10^{15}$ Jm^{-3}, obtained using NASA accurate measurement of the Pioneer Anomaly when Pioneer 10 was at 20 AU, see Appendix B.

7 Discussion

Eq. (2) assumes a spherical mass distribution for the mass of the Earth or Sun in the calculation of the gravitational energy density. It does not consider the possible influence of the Earth's oblate shape, which is known to affect orbiting spacecraft and could affect hyperbolic orbits.

Estimation has been done of the magnitude of the mass of Earth that deviates from spherical shape in order to calculate to what extent this can affect the gravitational energy density along the Flyby Anomaly trajectory. The calculation gives that the non spherical mass is of the order of less than 0.337% of the Earth mass. This amount influences the third term of the denominator in (9) and quantities derived from it. However, the subtraction or addition of this mass to the mass of Earth on the SQR term of (9) affects this term in less than the tenth significant figure. This estimate implies that the mass of Earth causing the gravitational quadrupole does not affect the calculations based on the Céspedes-Curé hypothesis.

The hypothesis also predicts that ranging measurements based on a constant value of c will be affected in the same manner as the anomalous speed measurements based on the Doppler data. Anomalous ranging is briefly mentioned by Anderson et al. [2]. However, no numerical data of this anomaly has been provided. Perhaps due to the small signal-to-noise ratio on the incoming ranging signal and a long integration time (typically minutes) that must be used for correlation purposes [21, page 7].

We calculate the speed of light at the International Space

Station to be

$$c' = 299798845.6 \, \text{ms}^{-1},$$

that is $6387.6 \, \text{ms}^{-1}$ higher than c on the Earth's surface, about 0.002% [31]. Ranging measurements based on a constant c that is lower than is predicted by this theory will be in slight error. And the error will be in the same manner as the anomalous speed measurements. The Cépedes-Curé hypothesis predicts the anomalous measurements of the Pioneer spacecraft without any adjustable parameter [27]. There are reports that that the Pioneer Anomaly was resolved as a thermal effect on papers by Rievers and Lammerzahl [15], Turyshev et al. [32] and Francisco et al. [33]. These reports do complex parameterized models of the thermal recoil to explain the anomaly.

We have reasons to doubt this explanation:

First. A detailed paper about the Pioneer Anomaly (55 pages in Phys. Rev. by Anderson et al. 2002) [21] clearly argues (see sections VIII. B, C and D, pages 32–35) that thermal recoil cannot account for the anomaly,

Second. Rievers and Lämmerzahl [15] do a very complex computational model of the spacecraft constructing all parts of the spacecraft internal and external in finite elements; assigning thermal, and radiative properties for each component, (absorption, reflection and emittance coefficients) in order to arrive at their resulting thermal radiation pressure.

Turyshev et al. [32] do a complex parameterized model for the thermal recoil force of the Pioneer spacecraft with several adjustable parameters. In particular the two adjustable parameters of Eq. (1) on page 2 predict the anomaly. However, any other parameters would negate the thermal origin of the anomaly.

Francisco et al. [33] use different modeling scenarios resulting in different acceleration values and choosing the 4th one with which a Monte Carlo modeling procedure is used to arrive at a value of the reported acceleration of the Pioneer 10 at an instant 26 years after launch.

All of these reports imply models with numerous adjustable parameters which could disprove the thermal origin of the anomaly.

Third. If the anomalous acceleration towards the sun depended on the thermal emission of heat from the RTG, Plutonium ^{238}Pu power sources, with a half life time of 87.74 years, the anomalous acceleration should decrease in time at the same rate, however, this is contrary to the almost flat long term behavior observed [21].

Forth. An anomaly similar to the Pioneer spacecraft was detected in Galileo spacecraft (see Section V. C, page 21) with a value of (acceleration) of $(8 \pm 3) \times 10^{-8} \, \text{cm/s}^2$, a value similar to that from Pioneer 10, with additional evidence based on ranging data, and in the Ulysses spacecraft (see Section V. D, page 21) Ulysses was subjected to an unmodelled acceleration towards the Sun of $(12 \pm 3) \times 10^{-8} \, \text{cm/s}^2$, in Anderson et al. [21]. Both spacecraft have completely different geometries and the thermal recoil theory is not applicable to them.

There are some unexplored fundamental aspects to the Céspedes-Curé hypothesis. The elementary relation (4) that is deduced for the relative speed of light c' measured on a space site relative to c on Earth, coupled to Einstein's relation for the rest mass $E = mc^2$ leads to an analytical relation that predicts Mach's principle, i.e. that mass and inertia depend on the far away stars and galaxies. Likewise, the Céspedes-Curé Hypothesis coupled to the electromagnetic expression for the speed of light, $c = 1/\sqrt{\varepsilon_0 \mu_0}$ leads to a direct relationship between the electromagnetic and gravitational forces.

8 Conclusions

The values shown in Table 2 indicate that the Flyby Anomaly can be accurately predicted by the theory presented in this work. This theory is capable of explaining qualitatively and quantitatively the anomaly, both, the measured positive, null and negative values. To calculate exact values of the anomaly of a spacecraft it is necessary to know the incoming and outgoing points where the spacecraft velocity was measured. The precise calculation of the Flyby Anomaly provides additional confirmation of the Céspedes-Curé hypothesis, that c the speed of light depends on the gravitational energy density of space as defined by (1) namely:

$$c' = \frac{k}{\sqrt{\rho'}} \, .$$

The evidence presented in this work for the Céspedes-Curé hypothesis has profound consequences in the current cosmology theories since it implies a revision of all astronomical measurements of velocity based on the Doppler, blue and red shifts, of stars and galaxies. These have importance in determination of matters such as the Hubble constant, the expansion of the universe, the flat rotation curve of galaxies (which gave birth to the theory of dark matter) and the extreme values of the redshifts of very far away galaxies (so called inflation) which gave birth to the theory of dark energy. These redshifts do not follow the linear relation proposed by Hubble but rather seem to imply an accelerated rate of expansion. The theories that follows from this hypothesis, the evidence and attempts to gather evidence for it and some of its consequences on current physics are explored in [18] and in the unpublished work mentioned above in [31].

Appendix A. Supporting data (Céspedes-Curé)

See Table 3: Data of starlight deflection measurements, reported by P. Merat [34] (δ in seconds of arc) at different distances from the Sun during total eclipses, used by J. Céspedes-Curé [19, see page 279], to calculate $\rho^* = 1.094291 \times 10^{15} \, \text{Jm}^{-3}$, the energy density of space due to far-away stars and galaxies.

Appendix B. Supporting data (Greaves)

Data used by E. D. Greaves in [26] for the arithmetic to calculate $\rho^* = 1.0838 \times 10^{15} \, \text{Jm}^{-3}$, the energy density of space

Table 3: Data of starlight deflection measurements, reported by P. Merat [34] (δ in seconds of arc) at different distances from the Sun during total eclipses, used by J. Céspedes-Curé [19, see page 279], to calculate $\rho^* = 1.094291 \times 10^{15}$ Jm^{-3}, the energy density of space due to far-away stars and galaxies.

Row	r (R_o Units)	$\delta \pm \Delta\delta$ (Merat)
1	2.09	1.02 ± 0.11
2	3.12	0.67 ± 0.08
3	4.02	0.58 ± 0.04
4	5.10	0.40 ± 0.07
5	6.06	0.41 ± 0.04
6	7.11	0.31 ± 0.04
7	7.84	0.24 ± 0.04
8	9.51	0.20 ± 0.06
9	11.60	0.16 ± 0.03

due to far-away stars and galaxies.

The calculation uses the following equations from [26]:

$$\text{Eq. (8)} \quad \rho^* = \frac{\rho_{\text{Sfar}} + \rho_{\text{Efar}} - n'^2(\rho_{S\,1AU} + \rho_E)}{n'^2 - 1}, \text{ and}$$

$$\text{Eq. (19)} \quad n' = 1 - \frac{E_D\,c}{2f_e\,G\left(\frac{M_S}{r_S^2} + \frac{M_E}{r_E^2}\right)},$$

where: (numerical values in SI units)

n', index of refraction of space at 20 AU (comes out to 0.999973567943846),

ρ^*, energy density of space due to far-away stars and galaxies,

E_D, a steady frequency drift of 5.99×10^{-9} Hz/s from the Pioneer 10 spacecraft [21, page 20],

$f_e = 2295$ MHz, the frequency used in the transmission to the pioneer spacecraft [21, page 15],

$c = 299792458.0$ m/s. Speed of light on Earth at surface,

$G = 6.67300 \times 10^{-11}$ m^{3}kg^{-1}s^{-2}, Newton's universal constant of gravitation,

$M_S = 1.98892 \times 10^{30}$ kg, mass of the Sun,

$M_E = 5.976 \times 10^{24}$ kg, mass of the Earth,

1 Astronomical Unit (AU) = 149 598 000 000 m.

The distances r_S and r_E are the distances from the spacecraft at 20 AU (20 AU from the Sun, 19 from Earth) to the center of the Sun and Earth respectively. To calculate Eq. (8) of [26] use is made of the energy density ρ_i given by Eq. (4) also of [26]:

$$\rho_i = \frac{GM_i^2}{8\pi r^4},$$

where r is the distance from the centre of the Sun or Earth to the point where the energy density is being calculated as follows:

For the Earth's surface: $r_E = 63781.40$ m, radius of Earth,

For the Sun at 1 AU: $r_S = 149598000000$ m,

For the Sun at 20 AU: Twenty times the previous value used to calculate ρ_{Sfar},

For the Earth at 20 AU: radius of earth + 19 times 149 598 000 000 m used to calculate ρ_{Efar}.

All values were calculated with Microsoft Office Excel 2003 which uses 15 significant digits of precision.

Acknowledgements

References

1. Anderson J. D., Campbell J. K. and Nieto M. M. The energy transfer process in planetary flybys. *New Astronomy*, 2007, v. 12 (5), 383–397. arXiv: astro-ph/0608087. DOI: 10.1016/j.newast.2006.11.004.

2. Anderson J. D., Campbell J. K., Ekelund J. E., Jordan E. and Jordan J. F. Anomalous Orbital-Energy Changes Observed during Spacecraft Flybys of Earth. *Phys. Rev. Letters*, 2008, v. 100, 091102, 1–4.

3. Jouannic B., Noomen R. and van den Ijssel J. A. A. The Flyby Anomaly: An Investigation into Potential Causes. Proceedings of the 25th International Symposium on Space Flight Dynamics ISSFD, Munich, Germany, 19–23 October 2015.

4. Lämmerzahl C., Preuss O., and Dittus H. Is the physics within the Solar system really understood. arXiv: gr-qc/0604052.

5. Acedo L. The Flyby Anomaly in an Extended Whitehead's Theory. *Galaxies*, 2015, v. 3 (3), 113–128. DOI: 10.3390/galaxies3030113.

6. Iorio L. A flyby anomaly for Juno? Not from standard physics. *Advances in Space Research*, 2014, v. 54 (11), 2441–2445.

7. Acedo L. The flyby anomaly: A case for strong gravitomagnetism? *Advances in Space Research*, 2014, v. 54 (4), 788–796.

8. Varieschi G. U. Kerr metric, geodesic motion, and flyby anomaly in fourth-order conformal gravity. *Gen. Relativ. Gravit.*, 2014, v. 46, 1741. DOI: 10.1007/s10714-014-1741-z

9. Acedo L. Kinematics effects of atmospheric friction in spacecraft flybys. arXiv: space-ph/1701.06939v1.

10. McCulloch M. E. Modeling the flyby anomalies using a modification of inertia. *Monthly Notices of the Royal Astronomical Society: Letters*, 2008, v. 389 (1), L57–L60.

11. Adler S. L. Can the flyby anomaly be attributed to earth-bound dark matter? arXiv: astro-ph/0805.2895v4.

12. Wilhelm K. and Dwivedi B. N. Anomalous Earth flybys of spacecraft. *Astrophys Space Sci.*, 2015, v. 358, 18. DOI: 10.1007/s10509-017-3205-x

13. Atchison J. A. and Peck M. A. Lorentz accelerations in the Earth flyby anomaly. *J. Guid. Control Dyn.*, 2010, v. 33, 1115–1122.

14. Acedo L. Anomalous accelerations in spacecraft flybys of the Earth. arXiv: astro-ph/1711.02875v2.

15. Rievers B. and Lämmerzahl C. High precision thermal modeling of complex systems with application to the Flyby and Pioneer Anomaly. *Annalen der Physik*, 2011, v. 523 (6), 439–449.

16. Cahill R. T. Resolving Spacecraft Earth-Flyby Anomalies with Measured Light Speed Anisotropy. *Progress in Physics*, 2008, v. 4 (3), 9–15. arXiv: gen-ph/0804.0039.

17. Greaves E. D. La hipótesis de Céspedes-Curé y el índice de refracción del espacio en un campo magnético. (The Céspedes-Curé hypothesis and the index of refraction in a magnetic field). *Acta Científica Venezolana*, 2015, v. 66 (4), 226–229.

18. Greaves E. D. Propiedades del espacio vacío. (Properties of empty space). Memorias del II Congreso de ABAE. September 18–22. 2017. `http://2cvte.abae.gob.ve/ejes.php?idiomas=es`. Retrieved 10 January 2019.

19. Céspedes-Curé J. Einstein on Trial or Metaphysical Principles of Natural Philosophy. 1st ed. Ramsey Laboratory, Inc, 2002. ISBN: 978-0971387300. `http://www.nuclear.fis.usb.ve/Cespedes-Cure-2002-Einstein-on-Trial-J.pdf`. Retrieved 10 January 2019.

20. Struve O. Elementary Astronomy. 1st ed. New York, Oxford University Press. 1959.

21. Anderson J. D., Laing Ph. A., Lau E. L., Liu A. S., Nieto M. M. and Turyshev S. G. Study of the anomalous acceleration of Pioneer 10 and 11. *Phys. Rev. D*, 2002, v. 65, 082004.

22. Einstein A. Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen. (About the principle of relativity and the consequences derived from it). *Jahrbuch für Radioaktivität und Elektronik*, 1907, v. 4, 411–462.

23. Einstein A. Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes. (About the influence of gravity on the propagation of light). *Annalen der Physik*, 1911, v. 35 (10), 898–906. DOI: 10.1002/andp.19113401005.

24. Dicke R. Gravitation without a Principle of Equivalence. *Reviews of Modern Physics*, 1957, v. 29 (3), 363–376. DOI: 10.1103/RevModPhys.29.363.

25. Magueijo J. New varying speed of light theories. *Reports on Progress in Physics*, 2003, v. 66 (11), 2025–2068. DOI: 10.1088/0034-4885/66/11/R04.

26. Greaves E. D. NASA's astonishing evidence that *c* is not constant: The Pioneer Anomaly. arXiv: gen-ph/0701130.

27. Greaves E. D. A Neo-Newtonian Explanation of the Pioneer Anomaly. *Rev. Mex. AA (Serie de Conferencias)*, 2009, v. 35, 23–24.

28. Halliday D. and Resnick R. Physics for Students of Science and Engineering, Part II. John Wiley & Sons, Inc, New York and London, 1960.

29. Bate R., Mueller D. and White J. Fundamentals of Astrodynamics. 1st ed. Dover Publications Inc, New York, 1971.

30. Kaplan M. Modern Spacecraft Dynamics and Control. 1st ed. John Wiley & Son Inc, New York, 1976, 287–289.

31. Greaves E. D. The index of refraction of quasi-empty space. Universidad Simón Bolívar, Caracas Venezuela. 2015. Unpublished. `http://www.nuclear.fis.usb.ve/fn/wp-content/uploads/2015/07/GREAVES-ED-Index-of-refraction-of-quasi-empty-space-V11.pdf`, Retrieved 19 April 2019.

32. Turyshev S. G., Toth V. T., Kinsella G., Lee S. C., Lok S. M. and Ellis J. Support for the Thermal Origin of the Pioneer Anomaly. *Phys. Rev. Letters*, 2012, v. 108 (24), 241101. arXiv: gr-qc/1204.2507. Bibcode: 2012PhRvL.108x1101T. DOI: 10.1103/PhysRevLett.108.241101.

33. Francisco F., Bertolami O., Gil P. J. S. and Páramos J. Modelling the reflective thermal contribution to the acceleration of the Pioneer spacecraft. arXiv: space-ph/1103.5222v2.

34. Merat P. Analysis of the optical data on the deflection of light in the vicinity of the solar limb. *GRG*, 1974, v. 5 (3), 757–764.

——————————

Gravity in the Microworld

Anatoly V. Belyakov

E-mail: belyakov.lih@gmail.com

A brief review article gives examples of using the physical model based on the mechanistic interpretation of J. Wheeler's geometrodynamics. The examples show the need to consider gravity in the microworld. The latter is based on the balance of magnetic and gravitational forces. The gravitational constant was used in calculating the masses of quarks, neutrinos, proton size, coupling constants, etc. A new deviation of 28 GeV in the physical experiments of CMS Collaboration was confirmed by calculations. The unusual value of s- quark and b- quark masses is explained.

1 Introduction

In the Standard Model of Fundamental Interactions (SM), gravitational forces are not taken into account. However, the model based on the geometrodynamics of John Wheeler (Wheeler John Archibald) has proved the need for introducing gravitational forces into the microworld.

In the mechanistic interpretation of J. A. Weheler's geometrodynamic, charged microparticles are singular points on a non-simply connected two-dimensional surface of our world, connected by a "wormhole" or a drain-source current line in an additional dimension, forming a closed contour. But "wormholes", by necessity and by virtue of physical analogy in their mechanistic interpretation, can only be vortex current tubes, where the charge is in the "coulombless" form proportional to the medium momentum along the vortex current tube, spin, respectively, to the angular momentum relative to the longitudinal contour axis, and the magnetic interaction between the conductors is similar to the forces acting between the current tubes [1].

In this model, the electron size with mass m_e and radius r_e is taken as a medium unit element, and then the contour mass becomes proportional to its length. It is this hidden mass and its motion that is responsible for gravity, charge, spin, and magnetic interaction in the microworld. The introduction of gravity into the microworld allows one to explain various micro-phenomena and in some cases to calculate some important parameters quite accurately, using only fundamental constants and an elementary mathematical apparatus.

2 On the structure of microparticles

Thus, microparticles are not point objects, but are likened to vortex formations in an ideal fluid, which can reside in two extreme forms — the vortex *on the surface* of radius r_x along the X-axis (let it be the analog of a fermion of the mass m_x) and the vortex thread *under the surface in depth* of radius r, of the angular velocity v, and of the length l_y, filling the current tube of the radius r_e along the Y-axis (let it be the analogue of a boson of the mass m_y).

In a real medium these structures oscillate, passing into each other (oscillation of oscillators), where fermions retain part of the bosonic mass, introducing a half spin. Note that bosonic masses cannot in principle be stable, like their physical counterparts — vortex formations in a continuous medium (if they do not lean on a phase boundary). The parameters of the vortex thread m_y, v, r, l_y for an arbitrary p^+-e^--contour were determined in dimensionless units of the electron mass m_e, its classical radius r_e, and the speed of light c [2]:

$$m_y = l_y = (an)^2, \tag{1}$$

$$v = \frac{c_0^{1/3}}{(an)^2}, \tag{2}$$

$$r = \frac{c_0^{2/3}}{(an)^4}, \tag{3}$$

where n is the main quantum number, a is the inverse fine structure constant, while c_0 is the dimensionless light velocity $c/[\text{m/sec}]$.

It is further shown particles themselves to be similar to the contour and have their own quantum numbers n_i, which determine, as it were, the zone of influence of these microparticles with the size $l_i = (an_i)^2$. For the proton and electron n_i are 0.3338 and 0.5777, respectively. A vortex tube of radius r_e is filled spirally with a vortex thread; therefore, with extreme "compression" and full filling, its length along the Y-axis is shortened proportional to $1/r$. In this case its compressed length $L_p = l_y r$ coincides numerically with the boson contour mass energy of units $m_e c^2$, and then it is true:

$$L_p = l_y r = m_y r = m_y v^2 = \frac{c_0^{2/3}}{(an)^2}. \tag{4}$$

It is obvious that an arbitrary boson mass in the mass-energy units will match of its own numerical value m_y only when the vortex tube ultimate excitation's case, wherein we have $r \to r_e$, $v \to c$, and $n_i \to 0.189$ (in experiments at high energie, for example). According to [2], the standard contour bosonic mass m_y is $c_0^{2/3} = 4.48 \times 10^5$ (in units of m_e), which approximately corresponds to the summary mass of W, Z-bosons. Therefore, it can be argued the vortex current tube to be form by three vortex threads rotating around

m_x	6.10×10^6	2090	1	$(4.4 \pm 0.1) \times 10^{-7}$
n_i	0.189	0.334	0.577	1.643
r_x	669	2090	6270	5.07×10^4
$m_k = L_p$	1.02	1.80	3.10	8.83
n	4.88	3.64	2.77	1.643
$l_y = m_y$	4.48×10^5	2.49×10^5	1.44×10^5	5.07×10^4

a common longitudinal axis. These threads are finite structures. They possess, by necessity, the right and left rotation; the last thread (it is evidently double one) possesses summary null rotation. They can be associated with the vector bosons W^+, W^-, Z^0.

This model assumes that a closed contour is created between charged particles in a region X (a p^+-e^--contour, for example); and only a temporary contour appears in a region Y, when a case of the weak interaction occurs (when a proton absorbs an electron, for example). The temporary contour then loses its charge (longitudinal momentum) and becomes a one-dimensional neutrino vortex tube, retaining spin. Since current tubes (i.e. field lines of some field) are treated as material objects, there are gravitational and magnetic interactions between them.

For a counter-currents closed contour the characteristic contour size l_k, which is the geometric mean of two linear quantities, is derived. This size is based on the balance of gravitational and magnetic forces written in the "Coulombless" form [2]. Applied to the X-axis l_k is:

$$l_k = (l_x r_x)^{1/2} = \left(\frac{z_{g_1} z_{g_2}}{z_{e_1} z_{e_2}}\right)^{1/2} (2\pi\gamma\rho_e)^{1/2} \times [\text{sec}], \qquad (5)$$

where z_{g_1}, z_{g_2}, z_{e_1}, z_{e_2}, r_x, l_x are gravitational masses and charges expressed through the mass and charge of the electron, the distance between the current tubes (charges) and theirs length, γ is the gravitational constant, while ρ_e is the electron density $m_e/r_e^3 = 4.07 \times 10^{13}$ kg/m³.

A vortex tube having a momentum equivalent to the electron charge was shown in [3] really to contain three single vortex threads (the calculated value is 2.973). These unidirectional vortex threads rotate about a longitudinal axis. Their peripheral speed v_0 is derived from the balance of magnetic and inertial (centrifugal) forces. In the case of unit charges, it is equal to:

$$v_0 = \frac{r_e}{(2\pi)^{1/2} \times [\text{sec}]} = 1.12 \times 10^{-15} \text{ m/sec}, \qquad (6)$$

and does not depend on the length of the vortex threads and the distance between them.

3   On the weak interaction

The proton has a complex structure, and quarks are in this model an active part of its mass, a kind of ring currents inside the proton, where in three local sections the medium velocity reaches critical parameters [2]. In the p^+-e^--contour, proton quarks are involved in the circulation, and their mass as z_{g_1} is included in equation (5) and depends on the contour size. For the weak interaction, the contour is limited only by its influence zone $l_i = (an_i)^2$. Setting $r_x = l_i$ and taking into account formulas (1–5), for the mass of quarks at unit charges, we obtained:

$$m_k = z_g = \frac{a n_i c_0^{1/3}}{2\pi\gamma\rho_e \times [\text{sec}^2]}. \qquad (7)$$

It should be noted that the quarks charges are integer ones inside the proton, and in the form of fractional quantities they are only projected onto the outer surface of the proton.

In the case of the weak interaction (electron absorption by the proton) the quark mass-energy is assumed to compar with the compressed bosonic contour mass-energy L_p in the Y-region, which, having lost a longitudinal momentum (charge), becomes the bosonic neutrino vortex tube [4]. This process is something similar to the charge and spin separation — a phenomenon registered in ultrathin conductors [5], which can be likened to a one-dimensional vortex current tube. Under this condition $m_k = L_p$, the quantum number n of Y-contour is calculated from formula (4), and the mass m_y (relative length) according to formula (1).

Table 1 that above shows the calculated parameters under various conditions of the weak interaction, i.e. for various distances between the proton and the electron, namely: the characteristic masses of fermions m_x, their own quantum numbers n_i, distances between charges r_x, quark masses m_k, boson tube quantum numbers and masses n and m_y.

The relationship between the fermion and bosonic masses was established in [2]. The most probable fermionic mass of neutrinos was determined in [4] under the additional condition of symmetry, when $r_x = l_y$ and $n = n_i$ (see Table 1); moreover, there are three more independent formulas containing the gravitational constant and giving actually that the result, equal to 4.4×10^{-7} (0.225 eV). It is not known whether neutrinos appear as a fermion at higher n; in these cases, their masses would be negligible, because they are inversely proportional to n^{14}. As for the structure of the neutrino, then, having no charge, it should have a closed shape. Apparently, the bosonic vortex tube, consisting a total of four vortex threads, is as a result organized into a pair of closed vortex

threads with left-right rotation and, conversely, with right-left rotation (with respect to the motion axis direction).

The minimum quark mass, as follows from the table, matches to the electron mass, and the most probable one (when neutrino is released) matches to the d-quark mass of 8.83 (4.8 MeV). The bosonic masses m_y are close to the masses of three, two (Higgs mass), one and one third of the W, Z particles masses. Although the last boson with a mass of 5.07×10^4 (26 GeV) has not yet been detected, events with close energies of about 28 GeV have already been recorded in the CMS Collaboration experiments [6].

These bosons are considered, on the one hand, to be truly fundamental particles, and on the other, to be pointlike virtual particles, moreover having enormous mass-energy. This fact is in no way compatible with the particles or atoms internal energy. They exist only about 10^{-25} seconds, although the duration of the weak interaction is $t > 10^{-12}$ seconds. The latter in this model is understandable, because t determines the time of a medium runing with speed v around the entire "extended" contour length. That is, given (1–4), we have:

$$ t = \frac{a^8 n^8}{c_0} \frac{r_e}{c}, \qquad (8) $$

that in the indicated range n gives $10^{-9} \dots 10^{-13}$ seconds, there is an interval corresponding to possible times of the weak interaction.

Given the inconsistencies in the W and Z bosons properties and based on the calculated masses m_y, these bosons (including the Higgs boson) are probably not fundamental particles, but rather the excited boson forms of neutrinos, which during high energy experiments acquired (or did not have time to lose) for a short time a longitudinal momentum (charge).

4 On the coupling constants

It was found [7] that the formula for the number of threads in a vortex tube, cubed, is the ratio of the inertia forces arising from the acceleration of the bosonic standard contour mass and acting towards the periphery, to the gravitational forces acting between fermionic masses of m_e at a distance r_e. The numerator is a constant, so this dependence is only determined by gravity, i.e. interacting masses and the distance between them

$$ n_i^3 = \frac{m_e c_0^{2/3} r_e / ((2\pi)^{1/2} \times [\sec^2])}{(2\pi)^{1/2} \gamma m_e^2 / r_e^2} = 26.25. \qquad (9) $$

This formula indicates the strength of bonds between the structural elements of microparticles (quarks) and, as it turns out, can serve as the equivalent of the coupling constant a_s for weak and strong interactions. Suppose that quarks are located in the corners of a regular triangle at a distance r_e. Then, taking into account the geometry of their interaction

and after calculating the constants, the formula (9) can for the general case be represented in a dimensionless form:

$$ a_s = 15.15 \, (r/m)^2. \qquad (10) $$

At low energies of interacting particles, affecting only the external structure of nucleons (small "depth" along Y), the peripheral inertia forces exceed the attractive forces, therefore quarks are weakly coupled to each other within a vortex tube of radius r_e, and they interact with quarks of nearby nucleons. At high energies (about 100 GeV, a great "depth" along Y) they reach within the proton itself vortex thread the minimum distance of $r \sin 60°$ (here r is calculated from (3)); in this case the mutual attraction forces keep the quarks in a bound state within the nucleon size. Then, with the quark minimum mass, $m_k = 1$, substituting $r = 1$ and $r = 0.0887$ in (10), we obtain: $a_s = 15.15$ and $a_s = 0.119$. These values coincide with the actual ones.

The validity of the above is also convincingly confirmed by the determination of the proton radius r_p provided that $a_s = 1$ and $m_k = 1$. Obviously, it is the vortex tube circumferential size and it is equal to $r/\sin 60°$. Revealing the constant in (10) and using the above formulas, we finally get:

$$ r_p = \left(\frac{8\pi\gamma\rho_e}{3^{1/2} c_0^{2/3}} \right)^{1/2} \times [\sec] = 0.297 \text{ or } 0.836 \text{ Fm}, \qquad (11) $$

which *exactly coincides* with the value obtained in recent experiments (0.833 femtometers, with an uncertainty of ±0.010 femtometers) [8].

In the weak interactions, bosonic vortex tubes take part in, but since their mass is high, the coupling constant for the weak interaction is very low (about 10^{-5}). With increasing interaction energy, vortex tubes are excited and their radius increases, and then this constant increases significantly. Thus, the coupling constant determines neither the nature of nuclear forces, nor the strength of interaction, but only indicates the strength of bonds within the complex structure of nucleons.

5 On the masses of s- and b- quarks

In [2] the total masses of the second and third generation quarks were approximately determined. But the masses of negative s- and b- quarks was in experiments found to be much smaller than the masses of their positive partners, and it can not be explained in SM. In this model the mass order of these quarks is at least reliably determined when using formula (5), derived from the balance of magnetic and gravitational forces. It was shown in [2] that any contour connecting charged particles can consider similar to a particle that is part of a larger contour, where the smaller contour mass is assumed to be a hypothetical fermion mass (a proton analog) for the larger one. Thus three generations of elementary particles are formed.

For the second generation (μ-contour), the proton analog is the mass of the standard contour $c_0^{2/3}$, for the third (τ-contour) one is the mass of the μ-contour, determined from the limiting conditions at $n_i = 0.189$ and equal to 6.10×10^6. Thus, for contours of subsequent orders it can be assumed of linear scale unit's increasing in proportion to the ratios of the μ-contour and τ-contour masses to the proton mass m_p. Since quarks masses is directly proportional to $l_x r_x$, i.e. to a linear parameter square, and inversely proportional to leptons masses, then, bearing in mind (5), we can write the relation:

$$\text{s- quark mass} \quad m_{ks} = \frac{m_k (c_0^{2/3}/m_p)^2}{m_\mu} = 222\, m_k \,,$$

$$\text{b- quark mass} \quad m_{kb} = \frac{m_k (6.10 \times 10^6/m_p)^2}{m_\tau} = 2450\, m_k \,,$$

where m_μ and m_τ are the μ- and τ- particles masses.

Consequently, the s- and b- quarks masses order is determined correctly: for the s- quark it is several hundred masses of the first generation quark, for b- quarks it is several thousand masses of the first generation quark.

6 Conclusion

Thus, the above examples show that gravity has a significant effect in the microworld, and the gravity constant should inevitably be included in the more accurate theories describing the microworld. Perhaps it is just this factor that may contribute to the further creation of the "theory of everything".

Submitted on April 15, 2020

References

1. Belyakov A.V. Charge of the electron, and the constants of radiation according to J. A. Wheeler's geometrodynamic model. *Progress in Physics*, 2010, v.4, 90–94.

2. Belyakov A.V. Macro-analogies and gravitation in the micro-world: further elaboration of Wheeler's model of geometrodynamics. *Progress in Physics*, 2012, v.2, 47–57.

3. Belyakov A.V. On Materiality and Dimensionality of the Space. Is There Some Unit of the Field? *Progress in Physics*, 2014, v.10, 203–206.

4. Belyakov A.V. Determination of the Neutrino Mass. *Progress in Physics*, 2016, v.12, 34–38.

5. Jompol Y., Ford C.J.B., Griffiths J.P., Farrer I., Jones G.A.C., Anderson D., Ritchie D.A., Silk T.W., and Schfield A.J. Probing Spin-Charge Separation in a Tomonaga-Luttinger Liquid. *Science*, 31 July 2009, v.325, no.5940, 597–601.

6. arXiv: 1808.01890 [hep-ex]. High Energy Physics — Experiment (hep-ex). CMS-HIG-16-017, CERN-EP-2018-204.

7. Belyakov A.V. Nuclear Power and the Structure of a Nucleus According to J. Wheeler's Geometrodynamic Concept. *Progress in Physics*, 2015, v.11, 89–98.

8. Bezginov N., Valdez T., Horbatsch M., Marsman A., Vutha A.C., Hessels E.A. A measurement of the atomic hydrogen Lamb shift and the proton charge radius. *Science*, 06 Sep. 2019, v.365, issue 6457, 1007–1012.

Application of the Theory of Hyperrandom Phenomena in the Search for Signs of the External Influence on Radioactive Decay and the Possibility of Quantitative Estimates

S. V. Adamenko[1], A. S. Kapshuk[1], V. E. Novikov[1], A. D. Skorbun[2*], S. N. Shpyl'ka[1], V. A. Yatsyshyn[1]

[1]Laboratory "Proton 23", 48-A, Chornovil Str., Vyshneve, Kyiv-Svyatoshyno district, Kyiv region, 08132, Ukraine.
[2]Institute for Safety Problems of Nuclear Power Plants of the NAS of Ukraine, 36a, Kirov Str., Chornobyl' 07270, Ukraine.
*Corresponding author E-mail: anskorbun@gmail.com

We have developed the relevant setup and studied a possibility of the influence on the radioactive decay by an external impulsive electromagnetic field. It is shown that such action can result not only in a change in the rate of decay (rate of counting of gamma-quanta), but also in a clear variation of the statistical properties of the series of successive measurements of the counting rate such as the appearance of periodicities and hyperrandom properties. It is found that the excitation of a system of radioactive nuclei induced by the external influence disappears approximately in 4–6 days.

1 Introduction

We will describe our attempt to find a possibility to affect parameters of the radioactive decay with the help of an impulsive electromagnetic field. As is known, at the radioactive decay, the number of decays per unit time is a random variable which is described by the Poisson distribution [1]. Hence, from the viewpoint of the statistical analysis, the problem of search for the signs of changes after some treatment of a radioactive specimen can be reformulated as a problem of changes in the statistical properties of samples which are the records of the results of measurements before and after the treatment.

It should be emphasized that we intend to seek the weak changes which can be only precursors of the changes seen by naked eye (hence, of those possessing a practical significance). From the viewpoint of the dominant theory, the rate of radioactive decay cannot be affected at all (see [2]). While experimentally determining the influence of some factor, the researchers try to find, as a rule, the changes in the counting rate at least on the level of the statistical effects. We pose the problem in a more general form: to seek the differences between samples which can or cannot be reduced to a change in the mean counting rate.

The sought signs can be the periodic variations in a counting rate or the appearance of irregular "splashes" of the intensity or other irregularities leading to that the series of measurements of the counting rate cease to be random in the sense of mathematical statistics. In this case, a change in the form of a distribution function (loss of the Poisson property) can be only one of the possible sought signs.

The radioactive decay can be considered as an example of the process (if the radioactive half-life is much more than the time of measurements), for which the long series of measurements of its parameters is considered to be stationary in the sense of mathematical statistics, i.e. its statistical parameters do not vary with time. For comparison, we can indicate examples of other natural processes without the property of stationarity such as the noise of the ocean, where ships move from time to time near a detector of noises. The problem of the analysis of such data was considered, for example, in [3, 4].

In the present work, we will analyze changes in the decay statistics for signals of the rate of counting of gamma-quanta from radioactive specimens after the action of an impulsive electromagnetic field onto them.

2 Data and methods of their analysis

We will examine a possibility to influence the process of radioactive decay by external impulsive electromagnetic field. The setup generating the electromagnetic impulses that act on a radioactive specimen will be called a driver for simplicity. In order to use the statistical methods of analysis, we need the long series of regular measurements of the rate of decay. Such series were recorded with the use of a dosimeter-radiometer "Pul's" aimed at the remote radiation control. The device was produced at the small joint-stock enterprise "Opyt", includes a detector on the basis of NaI(Tl), and allowed us to execute every-second measurements with the record of results into a memory unit.

We analyzed the results of measurements of a specimen treated with a driver during February–May in 2018 in the city of Chornobyl'. As a specimen, we took monazite sand, i.e. we measured and analyzed the summary signal (gamma-radiation) from decay products of ^{232}Th. First, before the treatment of the specimen with a driver, we carried out the measurements of the counting rate for several days. Later on, we compared those data with the results obtained after the action of a driver onto the specimen.

In the analysis of the statistical properties of the measured signals, we used the statistical theory of hyperrandom phenomena [5]. This theory is based on the hypothesis that the results of measurements of natural processes are not independent and identically distributed. Hence, they do not obey the

basic preconditions for the application of well-known methods of mathematical statistics. In other words, the basic assertion of the theory of hyperrandomness consists in that the process under study can undergo the action of external influences, which induces, respectively, changes in the statistics of a signal. This is manifested in the loss of the statistical stability by data, i.e. the results of measurements become dependent on the time. However, it can turn out that very long series of measurements should be made for such changes to be revealed.

The main distinction of the hyperrandom data from the standard random numbers which are independent and identically distributed consists in that the variance of the former does not decrease, as the number of measurements increases (increase in the size of a sample). On the contrary, starting from some number of measurements, the variance of hyperrandom data increases [3, 4]. Such effect can be a consequence of the tendency to a change in the mean, the autocorrelated function, *etc* during the measurement. (We emphasize once more that the similar changes in a sample can have the statistical character and can be invisible for naked eye.)

The formulas for the analysis of hyperrandom data can be found in [5, 6]. We now indicate only the principle of such analysis. Let us have the sample of the results of measurements X with size N: $X = (x_1, x_2, \ldots x_N)$ is a regular temporal series of the results of measurements. We accentuate that the series is ordered in the meaning that the elements of the series should not be permuted. We are interested in the dependence of its parameters on the size of a sample, i.e. on the time. For this purpose, we calculate the accumulated means, i.e. the means for the first two, three, *etc* elements of the input series. As a result, we get a new first-order series of data in the form of accumulated means $Y^{(1)} = (Y_1, Y_2, \ldots Y_N)$, where $Y_n = \frac{1}{n} \sum_{i=1}^{n} x_i (n = \overline{1, N})$, with its mean $\overline{m}_{Y_N} = \frac{1}{N} \sum_{n=1}^{N} Y_n$. Then we can repeat the procedure and form the series of higher orders $Y^{(2)}$, $Y^{(3)}$, *etc*.

The object of our analysis is the function, being the unbiased variance of fluctuations from the accumulated mean, $\overline{D}_{Y_N} = \frac{1}{N-1} \sum_{n=1}^{N} (Y_n - \overline{m}_{Y_N})^2$.

As the quantitative measure of one of the hyperrandom properties, specifically, the statistical instability of a series of data, we take the coefficient γ_N characterizing the absolute level of statistical instability: $\gamma_N = \frac{M[\overline{D}_{Y_N}]}{ND_{Y_N}}$, where $M[*]$ is the operator of mathematical expectation.

To have a possibility to compare different samples with one another, the units of statistical instability are introduced in the theory. For the coefficient γ_N, the role of a unit of statistical instability of measurements is played by the quantity γ_{0N} which corresponds to the noncorrelated series of readings with constant variance $D_{x_n} = D_x$ and zero mathematical expectation at a fixed value of N. The coefficient γ_{0N} is given by the formula

$$\gamma_{0N} = \frac{N+1}{(N-1)N} C_N - \frac{2}{N-1}, \text{ where } C_N = \sum_{n=1}^{N} \frac{1}{n}.$$

Using the unit of measurements γ_{0N}, we introduce the ratio $h_N = \frac{\gamma_N}{\gamma_{0N}}$, i.e. the coefficient characterizing the absolute level of statistical instability in units of γ_{0N}. These coefficients are dimensionless. The degree of hyperrandomness h_N of the analyzed data will be considered in what follows.

We note that though the hyperrandom properties of our data are manifested undoubtedly (see below), the derivation of the quantitative estimates of the degree of hyperrandomness is not a simple matter. We clarify this point by the example. Let us deal with a really random stationary process, so that its signal has no signs of the hyperrandomness. At some time moment, let a quite short external influence arise (the duration of the external action is assumed to be much less than the time of observations). It causes an increase in the mean and, respectively, to the appearance of the hyperrandomness. After some time period, the signal again becomes random and stationary.

Hence, the sample as a temporal signal can be partitioned into three parts. The midsection is hyperrandom, and the beginning and the end are normal stationary signals. In this situation, the sample has, on the whole, hyperrandom properties. But the results of calculations for each of the three parts separately will give different results.

In real situations, the information about the very fact of the external influence (treatment by a driver) can be unknown. Hence, we should consider the problem of determination of changes in the statistics of a series, the problem of analysis of the dynamics of those changes in time, and the problem of searching for the time, when the driver acts. If, for example, the aftereffect is present and varies in time, we can say nothing about the time moment of the transition of the sample into the third part, even if we are based on the analysis of the whole series. Moreover, the very fact of such transitions should be studied. We reformulate this problem as follows: Are there some regularities of changes in the hyperrandomness indicating the action of a driver and can we determine, for example, the characteristic time of relaxation of the "hyperrandomness state" arisen due to the action of a driver?

In view of the above discussion, we need to analyze the separate parts of samples with the purpose to find the distinctions between them and to establish the optimum size of such subsamples. To make it, we chose a "window" of a definite size, i.e. we set the size of a subsample. With such window, we scan the whole series of measurements. For each "window", we calculated the necessary parameters.

3 Results and discussion

As was indicated, the hyperrandomness by its nature arises at a change in time of some parameters of the process such
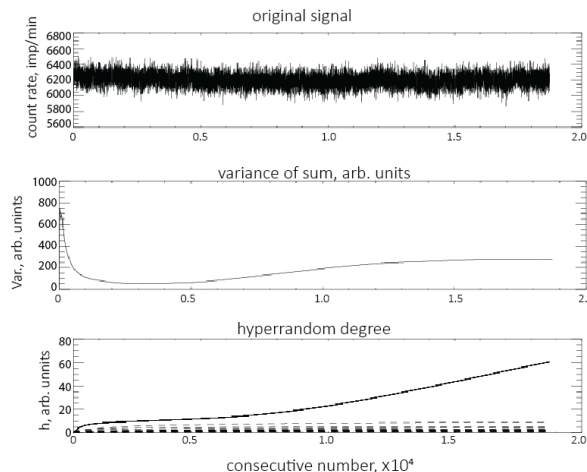
Fig. 1: Analysis for hyperrandomness of a series of measurements during 13 days from 14.02.2018 to 26.02.2018 (prior to the treatment). There is the sign of the hyperrandomness, which is revealed as an increase in h_N after approximately 5000–8000 min of measurements.

as, in particular, the solitary short-time splashes. The statistical characteristics of a series calculated before the splash can be changed after it. If the duration of the action of a driver is from several minutes up to several hours, it can be considered a short-time influence against the background of measurements during several days.

One of the tasks of the present work is the search for the time of relaxation of a signal after the action of a driver, which is reduced to the analysis of short segments of the entire series. In Fig. 1, we present the results of a test for the hyperrandomness. We took a sufficiently long-time (13 days) series of measurements before the action of a driver in order to estimate the order of long-time changes.

In each of the figures below, the upper plot is the input series of data; the middle plot presents a variation in time of the accumulated variance; and the lower plot shows the parameter h_N which characterizes the degree of hyperrandomness. The results of calculation of the hyperrandomness parameter are accompanied by the analysis of whether such result can be formed accidentally. It is a reasonable question, because we analyze the series of random numbers. For this purpose, we generated a computer-created sample of random numbers with the same parameters (mean and variance), as those of the experimental series. For such model sample with the same programs, we made analysis for hyperrandomness. This procedure was repeated several times for the sake of reliability, and the results were drawn on one figure. In the presence of a noticeable hyperrandomness, the experimental curve must be outside the zone, where the curves for model samples are placed. This zone for the model series of random numbers is shown in the lower plot by dotted lines.

As is seen in Fig. 1, the series manifests some hyperran-

domness during 13 days before the treatment. It starts to reveal itself after approximately 5–6 days of measurements.

Then, on 27.02.2018, we executed the treatment of the specimen with a driver (impulsive electromagnetic field).

In Fig. 2, we present the results of analysis for the hyperrandomness of a series of measurements before and after the action of a driver. We recall that our purposes are to register the time of a manifestation of the action of a driver and to determine the temporal changes of the signs of such action. We analyzed the subsamples 4 days in duration. In other words, we analyzed a part of the series 4 days in duration, then the "window" was shifted by one day, and so on. Hence, the subsamples were overlapped during 3 days in order to more or less reliably notice the times of changes in the degree of hyperrandomness.

In view of Fig. 2, we can formulate the following main results:

1. After the action of a driver, the rate of counting of gamma-quanta somewhat increased.

2. In the analyzed series, the hyperrandomness was not observed practically for 4 days (accepted size of a scanning "window") before the treatment: the variance decreased, as the size of a sample increased.

3. After the action of a driver on 27.02.2018, we observe a sharp increase in the hyperrandomness. The variance starts to grow already approximately in 1200 min (20 h).

4. This effect of hyperrandomness practically disappeared on 04.03.2018 (in 4–5 days) to the level of noises.

4 Conclusions

1. We have revealed that, under the action of electromagnetic impulses, the statistics of the radioactive decay is changed.

2. It is found that, after the action of a driver, the process of decay became hyperrandom. This means that its characteristic such as the accumulated variance increases in time, rather than decreases. In turn, this means that the process of decay stops to be stationary.

3. This forced nonstationarity was observed during approximately 5 days. Then the process of decay returns to the stationary mode (experimental curve in Fig. 2f is located in the zone of random values).

4. Such time of existence of the aftereffect (tens of hours), which is much more than the characteristic time of the evolution of a separate nucleus, is, most probably, the experimental confirmation of the theories (see [7–10]) that assert that the radioactive decay is a collective process in the system of correlated nuclei. From this position, we may assert that the quantitative estimates of the process of relaxation of a system of nuclei are made
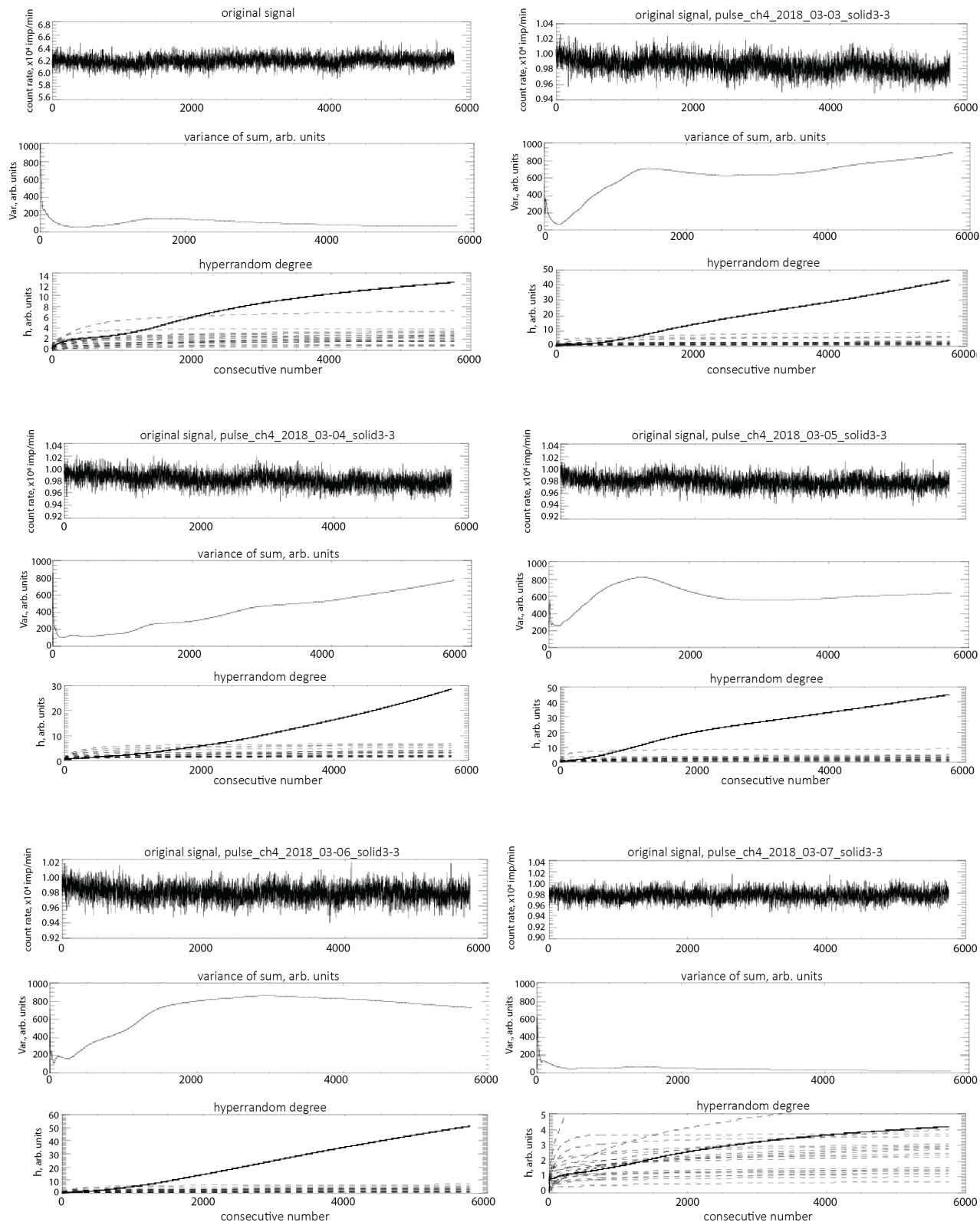
Fig. 2: Analysis for hyperrandomness: the series of successive overlapping subsamples from 28.02.2018 to 03.03.2018. The measurement at once after the treatment which occurred 27.02.2018. After the action of a driver, the hyperrandomness appeared: an increase in the variance and in *h* is clearly seen. On the fifth day, the hyperrandomness drops to the level of random noises.

for the first time. The determined time of the relaxation has the order of hours.

References

1. Gol'danskii V. I., Kutsenko A. V., Podgoretskii M. I. Statistics of Readings at the Registration of Nuclear Particles. GIFML (State Publishing House of Physical and Mathematical Literature), Moscow, 1959, p. 411.

2. Gareev F. A., Zhidkova I. E., Ratis Yu. L. Influence of the excitation and ionization of atoms in the rate of nuclear processes at low energies. Preprint R4-2004-68, JINR, Dubna, 2004, p, 54.

3. Gorban' I. I., Yaroshchuk I. O. Study of the statistical stability of oscillations of the temperature and the sound velocity in the ocean. Proceedings of the Acoustic Symposium "Consonance-2011", Institute of Hydromechanics of the NAS of Ukraine, Kiev, 2011, pp. 99–104.

4. El'yasberg P. E. Measurement Information: How Much Is It Needed? How Is It Processed? Nauka, Moscow, 1983, p, 208.

5. Gorban' I. I. Theory of Hyperrandom Phenomena. Institute of Problems of Mathematical Machines and Systems of the NAS of Ukraine, Kiev, 2007, p. 184. http://ifsc.ualr.edu/jdberleant/intprob/.

6. Gorban' I. I. Statistical instability of physical processes. *Radioelectronics*, 2011, v. 54 (9), 40–52.

7. Torchigin V. P., Torchigin A. V. Features of ball lightning stability. *Eur Phys J. D*, 2005, v. 32, 383—389. Torchigin V. P., Torchigin A. V. Ball Lightning (from the Improbable to the Obvious). MTsNMO, Moscow, 2009, p. 148.

8. Adamenko S. V., Selleri F., van der Merwe A. Controlled Nucleosynthesis. Breakthroughs in Experiment and Theory. Springer, Berlin, 2007.

9. Adamenko S. V., Vysotskii V. I. Correlated states of interacting particles and the problem of transparency of a Coulomb barrier at a low energy in nonstationary systems. *Zh. Tekh. Fiz.*, 2010, v. 80 (5), 23–31.

10. Adamenko S. V., Vysotskii V. I. Peculiarities of the formation and application of correlated states in nonstationary systems at a low energy of interacting particles. *ZhETF*, 2012, v. 141 (2), 276–287.

––––––––––––––––––

Periodic Phenomena in the Rate of Radioactive Decay
Under the Action of an Electromagnetic Field

S. V. Adamenko[1], A. S. Kapshuk[1], V. E. Novikov[1], A. D. Skorbun[2], S. N. Shpyl'ka[1], V. A. Yatsyshyn[1]

[1]Laboratory "Proton 23", 48-A, Chornovil Str., Vyshneve, Kyiv-Svyatoshyno district, Kyiv region, 08132, Ukraine.
[2]Institute for Safety Problems of Nuclear Power Plants of the NAS of Ukraine, 36a, Kirov Str., Chornobyl' 07270, Ukraine.
*Corresponding author E-mail: anskorbun@gmail.com

We have determined a mode of treatment of a radioactive material (^{232}Th and daughter products in a colloid solution of monazite sand) with a sequence of short impulses of an electromagnetic field which results in a change in the intensity (counting rate) of gamma-radiation. The value of changes in the intensity as a general trend is approximately 1.8 % for a period of about 1 month. In addition to the changes in the intensity, we observed changes in the statistics of the radioactive decay. In the long-term signal of every-second regular measurements of the counting rate for daughter products of the decay of ^{232}Th obtained after the treatment of the specimen with an impulsive electromagnetic field, we have found the periodic components among which the periods of 0.5, 1, and 6.6 days are distinguished most clearly.

1 Statement of the problem

The regular periodic changes in the intensity of signals of the radioactive decay, as well as the sporadic splashes of the counting rate, were observed many times (see [1]). They include the seasonal changes with periods of 1 year, 1 month, and 1 day which can be obviously related to astronomical phenomena. While analyzing the periodicity of the 1-day radioactive decay, it is necessary to separate the studies of the changes in the intensity of signals from radon and beryllium in the near-Earth layer of the atmosphere, which are caused mainly by geophysical or meteorological factors and are omitted in the present work, from the studies of the radioactive decay under controlled conditions (see [2–4] and references therein), where the manifestations of the variability of a signal are assigned to changes namely in the rate of radioactive decay. The presence of changes in the rate of decay can be related, in our opinion, to such fundamental causes as cosmophysical factors. Separately, we mention the works by S. E. Shnoll [5], where the regular changes in the statistics of separate parts of the series of measured data on the rate of radioactive decay, namely, the changes in a form of the distribution function, were observed. It is worth to note the fundamental cycle of works executed by opponents of this idea [6–8], where the special studies of this question revealed no existence of seasonal changes in the half-life period.

But earlier, the existence of that or other regularities of the radioactive decay was discussed without any experimental interference at laboratories. Moreover, the possibility of the influence of external physical factors, being outside the nuclear scale of energies, on the rate of radioactive decay was considered impossible since Rutherford's times [1, 9, 10]. Hence, the studies in this direction have a fundamental meaning, because they would prejudice the basic assertions of the theory that, first, all events of a radioactive decay are mutually in-

dependent, and, second, the internal processes in a nucleus which define the processes of decay can be affected only by the fluxes of particles and quanta with energies of the order of those of nuclear transitions from kiloelectronvolts to megaelectronvolts. However, we mention a well-known exclusion, isotope ^{229}Th, whose excitation energy is about 1 eV.

In this work, we present the results of laboratory studies of the influence of an external electromagnetic field with sufficiently low intensity on the rate of counting of gamma-quanta from a radioactive specimen and will show that such influence is possible.

Especially, we note that, though the revealed changes in the intensity of a signal after the action of an external factor for the period of observations up to 40 days are rather small, the changes in the statistical properties of the obtained regular series of measurements are obvious and objective.

2 Materials and methods

The setup realizing the action on a specimen (for simplicity, we call it a driver) is a system of coils with special structure aimed at the creation of an impulsive electromagnetic field. The duration of impulses is 1–10 nsec. The treatment was carried on for 10–30 min. The power of the setup is about 25 W. The energy of impulses is at most 2.5 J.

The scheme of the experiment is as follows. First, we carried on the control measurements for some time. After the action of a driver, the specimen was returned to a counter, and we measured the radiation from the specimen for several days. Such procedure can be repeated several times. In this case, we studied the intensity of the summary gamma-radiation from a specimen of natural monazite sand (mineral with ^{232}Th) for the period from 20.12.2017 to 15.01.2018. The results are obtained in a laboratory, i.e. under controlled conditions. During the indicated period of measurements, the driver did not act on the specimen, i.e. the presented results

are a manifestation of the aftereffect.

To gauge the counting rate, we used a counter of gamma-quanta such as a dosimeter "Pul's" for the remote radiation control. It was produced at the small joint-stock enterprise "Opyt" and includes a detector on the basis of NaI(Tl). The construction of the counter itself contains no lead-based protection. The counter can operate in the automatic mode and can write the result of measurements in the memory every second. During the measurements, the counter with a specimen have no special protection or can be placed inside a lead cylinder. The latter was open from one end, was about 30 cm in length, and has walls 10 cm in thickness. The measurements were carried out on different specimens, in different modes of action of a driver, many times, under the lead protection, and without it. On the whole, the results were invariable, i.e. the below-described effects did not disappear. The described conditions of measurements are given for the concreteness. The measurements were performed in heated premises. The changes on the temperature were in the interval 17–22 °C, but they were not regular with daily period.

In the room, where the measurements were performed, the background was much lower than the level of signals. For the indicated period, we have got a series of every-second measurements of the intensity of gamma-radiation with interruptions for the time, when the treatment of a specimen was executed.

Since the purpose of the present work is the search for the periodicity of a signal, we applied the wavelet-analysis using Gauss–Morlet wavelets [11, 12].

In Fig. 1, we show the signal from the untreated specimen in the form of a noisy path and its wavelet-expansion as a two-dimensional pattern of disordered spots. If some periodic regularities of the type of modulation by a sinusoid are present in the signal, the pattern of coefficients of the Gauss wavelet-expansion will contain the series of spots regularly arranged along the horizontal. The distance between such regular spots along the horizontal is equal to the half-period in units of the horizontal axis.

In view of the low rate of counting, while seeking the periodicities with a period of 1 h and more, we transformed the input series of data into a series of measurements for each 10 minutes (sum of sequential values for each 600 sec without the overlapping of intervals, where the number of counts is calculated).

3 Results

The analyzed signal itself after the action of a driver and the result of its wavelet-expansion are given in Fig. 2. The upper plot is the series of data which should be analyzed. Below, the two-dimensional pattern is the representation of coefficients of the wavelet-expansion (result of a wavelet-transformation of some series of data is the two-dimensional matrix of coefficients of the expansion). On the horizontal axis of the
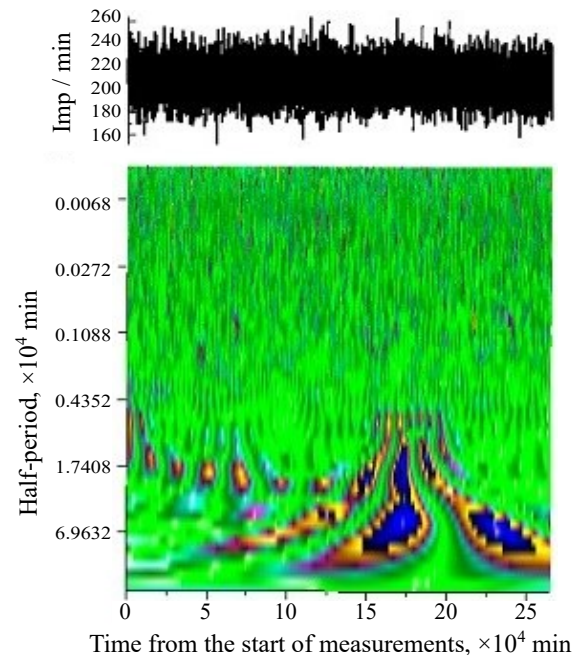

Fig. 1: Wavelet-expansion of a signal of the intensity of gamma-radiation from a specimen of monazite sand without the treatment with a driver. The upper plot is a signal; below, the two-dimensional pattern is the representation of the matrix of coefficients of the wavelet-expansion of this signal.

pattern, we indicate the number of a measurement which is proportional to the time from the start of measurements.

We note that the measurements are regular, and, in this case, each measurement corresponds to a time interval of 10 min. Therefore, for example, the number 1000 on the horizontal axis corresponds to a time moment of 10 000 min from the start of measurements. Along the vertical axis of the two-dimensional pattern, we give the half-period of a signal in units of the horizontal axis. We see clearly several horizontal rows of "spots," the distances between which are equal to the half-period (by assuming the modulation by a sinusoid).

In the table near the wavelet-expansion pattern, we show, as an example, several distances between spots for the rows indicated by arrows directly from the two-dimensional pattern in Fig. 2. This allows us to draw conclusion about the uncertainty of those estimates.

In addition to the appearance of a periodic daily modulation on the wavelet-expansion patterns, we observe that the signal itself looks as a uniform noisy path. We see also a tooth ripple of the daily variation and a small asymptotic decline.

4 Discussion of results and conclusions

The obtained array of results about the dynamics of the rate of counting of gamma-quanta after the treatment of a specimen with a driver testifies indisputably to the presence of the
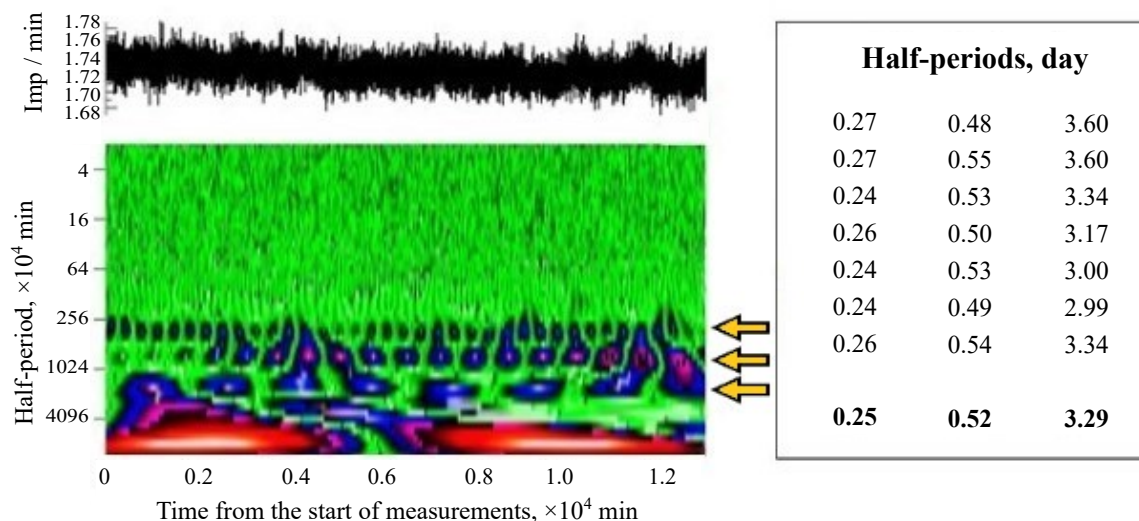
Fig. 2: Wavelet-expansion of a signal of the intensity of gamma-radiation of a specimen with Th after the treatment with a driver. The upper plot is the signal; below, the two-dimensional pattern is the representation of the matrix of coefficients of the wavelet-expansion of this signal. Arrows indicate the periodic series. The signal is the number of impulses for 1 min. The table on the right shows some examples of the estimates of distances between spots. The lower row gives the mean values.

external influence on the process of radioactive decay. This fact cannot be a result of erroneous measurements or, especially, improper analysis: the effect is not observed prior to the treatment and is seen after it.

Because the described result, i.e. the appearance of a periodicity in the noisy signal, is reliably established experimentally, we may ask: why was no effect observed earlier? The possible causes are as follows:

1. The effect is reliably registered only in definite operational modes of a driver. To observe it, one needs to perform special long-term experiments, since the effect can reveal itself in several tens of days after the action of a driver.

2. In the whole array of experimental data, the effect of a variation of the trend of the rate of counting is not strong (several percents) and is variable in time. Therefore, without regular long-term measurements, the effect can be considered as a noise or the uncertainty of a procedure of measurements.

Thus, we showed experimentally the presence of many periodicities with noninteger ratios in measured sequences of the rates of counting. The logical consequence of these results and of the up-to-date model representations about the fractal dynamics of intranuclear clusters [13–15] and about the nuclear structure [13–18] is the conclusion that sufficiently weak electromagnetic signals can excite the dynamics of intranuclear clusters, and, hence, it is possible to observe a manifestation of this changed dynamics in the probabilities of nuclear processes.

References

1. Parkhomov A. G. Deviations from Beta Radioactivity Exponential Drop. *Journal of Modern Physics*, 2011, v. 2, 1310–1317.

2. Sturrock P. A., Steinitz G., Fischbach E. Analysis of Radon-Chain Decay Measurements: Evidence of Solar Influences and Inferences Concerning Solar Internal Structure and the Role of Neutrinos. arXiv: astro-ph/1705.03010.

3. Steinitz G., Kotlarsky P., and Piatibratova O. Observations of the relationship between directionality and decay rate of radon in a confined experiment. *Eur. Phys. J. Special Topics*, 2015, v. 224, 731–740.

4. Jenkins J. H., Fischbach E., Buncher J. B., Gruenwald J. T., Krause D. E., Mattes J. J. Evidence for Correlations Between Nuclear Decay Rates and Earth-Sun Distance. arXiv: astro-ph/0808.3283v1.

5. Shnoll S. E. Cosmophysical Factors in Stochastic Processes. American Research Press, Rehoboth, NM, USA, 2012, p, 430.

6. Pomme S. *et al*. On decay constants and orbital distance to the sun – part II: beta minus decay. *Metrologia*, 2017, v. 54, 19–35.

7. Pomme S. *et al*. On decay constants and orbital distance to the Sun – part I: alpha decay. *Metrologia*, 2017, v. 54, 1–18.

8. Pomme S. *et al*. On decay constants and orbital distance to the Sun – part III: beta plus and electron capture decay. *Metrologia*, 2017, v. 54, 36–50.

9. Glasstone S. Sourcebook on Atomic Energy, Krieger, Malabar, FL, 1979.

10. Gareev F. A., Zhidkova I. E., Ratis Yu. L. Influence of the excitation and ionization of atoms in the rate of nuclear processes at low energies. Preprint R4-2004-68, JINR, Dubna, 2004, p. 54.

11. Astaf'eva N. M. Wavelet-analysis: foundations of the theory and applications. *Uspekhi Fiz. Nauk.*, 1996, v. 166 (11), 1145–1170.

12. Torrence C. A practical guide to wavelet analysis. *Bulletin of the American meteorological society*, 1998, v. 79 (1), 61–78.

13. Adamenko S. V., Selleri F., van der Merwe A. Controlled Nucleosynthesis. Breakthroughs in Experiment and Theory. Springer, Berlin, 2007.

14. Adamenko S., Bolotov V., Novikov V., Yatsyshin V. Control of multiscale systems with constraints 4. Control of the evolution of nuclear systems on the basis of the principle of dynamical harmonization. *Interdisciplinary Studies of Complex Systems*, 2013, No. 3, 35–95.

15. Greiner W. Exotic nuclei: from super-heavies to hyper and antimatter. *Yadern. Fiz.*, 2001, v. 64 (6), 1191–1197.

16. Maruyama T., Niita K., Oyamatsu K., Chiba S., and Iwamoto A. Quantum molecular dynamics approach to the nuclear matter below the saturation density. *Phys. Rev. C*, 1998, v. 57 (2), 655.

17. Dechargeé J., Berger J.-F., Girod M., Dietrich K. Bubble and semi-bubbles as a new kind of superheavy nuclei. *Nuclear Physics A*, 2003, v. 716, 55–86.

18. Berezovoj V. P., Bolotin Yu. L., Yanovsky V. V. *et al.* Stochastic Rezonance in nuclear fission. *Problems of Atomic Science and Technology*, 2001, No. 6, 226–229.

––––––––––––––––––––

Postal Address

Department of Mathematics and Science,
University of New Mexico,
705 Gurley Ave., Gallup, NM 87301, USA

This journal is powered by LaTeX

CONTENTS

Information for Authors

Progress in Physics has been created for rapid publications on advanced studies in theoretical and experimental physics, including related themes from mathematics and astronomy. All submitted papers should be professional, in good English, containing a brief review of a problem and obtained results.

All submissions should be designed in LaTeX format using *Progress in Physics* template. This template can be downloaded from *Progress in Physics* home page http://www.ptep-online.com

Preliminary, authors may submit papers in PDF format. If the paper is accepted, authors can manage LaTeXtyping. Do not send MS Word documents, please: we do not use this software, so unable to read this file format. Incorrectly formatted papers (i.e. not LaTeXwith the template) will not be accepted for publication. Those authors who are unable to prepare their submissions in LaTeXformat can apply to a third-party payable service for LaTeX typing. Our personnel work voluntarily. Authors must assist by conforming to this policy, to make the publication process as easy and fast as possible.

Abstract and the necessary information about author(s) should be included into the papers. To submit a paper, mail the file(s) to the Editor-in-Chief.

All submitted papers should be as brief as possible. Short articles are preferable. Large papers can also be considered. Letters related to the publications in the journal or to the events among the science community can be applied to the section *Letters to Progress in Physics*.

All that has been accepted for the online issue of *Progress in Physics* is printed in the paper version of the journal. To order printed issues, contact the Editors.

Authors retain their rights to use their papers published in *Progress in Physics* as a whole or any part of it in any other publications and in any way they see fit. This copyright agreement shall remain valid even if the authors transfer copyright of their published papers to another party.

Electronic copies of all papers published in *Progress in Physics* are available for free download, copying, and re-distribution, according to the copyright agreement printed on the titlepage of each issue of the journal. This copyright agreement follows the *Budapest Open Initiative* and the *Creative Commons Attribution-Noncommercial-No Derivative Works 2.5 License* declaring that electronic copies of such books and journals should always be accessed for reading, download, and copying for any person, and free of charge.

Consideration and review process does not require any payment from the side of the submitters. Nevertheless the authors of accepted papers are requested to pay the page charges. *Progress in Physics* is a non-profit/academic journal: money collected from the authors cover the cost of printing and distribution of the annual volumes of the journal along the major academic/university libraries of the world. (Look for the current author fee in the online version of *Progress in Physics*.)

————————

Fundamental Geometrodynamic Justification of Gravitomagnetism (I)

G. G. Nyambuya

National University of Science and Technology, Faculty of Applied Sciences – Department of Applied Physics,
Fundamental Theoretical and Astrophysics Group, P. O. Box 939, Ascot, Bulawayo, Republic of Zimbabwe.
E-mail: physicist.ggn@gmail.com

At a most fundamental level, gravitomagnetism is generally assumed to emerge from the General Theory of Relativity (GTR) as a first order approximation and not as an exact physical phenomenon. This is despite the fact that one can justify its existence from the Law of Conservation of Mass-Energy-Momentum in much the same manner one can justify Maxwell's Theory of Electrodynamics. The major reason for this is that in the widely accepted GTR, Einstein cast gravitation as a geometric phenomenon to be understood from the vantage point of the dynamics of the metric of spacetime. In the literature, nowhere has it been demonstrated that one can harness the Maxwell Equations applicable to the case of gravitation – i.e. equations that describe the gravitational phenomenon as having a magnetic-like component just as happens in Maxwellian Electrodynamics. Herein, we show that – under certain acceptable conditions where Weyl's conformal scalar [1] is assumed to be a new kind of pseudo-scalar and the metric of spacetime is decomposed as $g_{\mu\nu} = \mathcal{A}_\mu \mathcal{A}_\nu$ so that it is a direct product of the components of a four-vector $\mathcal{A}_\mu$ – gravitomagnetism can be given an exact description from within Weyl's beautiful but supposedly failed geometry.

My work always tried to unite the Truth with the Beautiful, but when I had to choose one or the other, I usually chose the Beautiful.

Herman Klaus Hugo Weyl (1885-1955)

1 Introduction

Exactly 102 years ago, the great, brilliant and esoteric German mathematician *cum* mathematical physicist and philosopher – Herman Klaus Hugo Weyl (1885-1955) – astounded the world of Physics with the first ever unified field theory of gravitation and electromagnetism. At the time, gravitation and electromagnetism were the only known forces of Nature, hence, from the viewpoint of the collective wisdom of the day, Weyl's [1] theory was seen as a unified field theory of all the forces of Nature. Since Weyl's [1] maiden efforts, unification of the gravitational phenomenon with the other forces of Nature has remained as one of the greatest – if not the greatest – and most outstanding problem in all of physics today. This endeavour of unification of all the forces of Nature first conducted by Weyl [1], became Albert Einstein's (1879-1955) final quest in the last 30 years of his brilliant and eventful life.

Since it is a widely accepted position, it perhaps is only fair for us to say at this very point, that – overall – while he failed in his titanic 30-year long quest and battle with the problem of an all-encompassing unified field theory of all the forces of Nature, Einstein made serious meaningful contributions to this seemingly elusive grand dream of a *Final Theory* that ties together all the known forces of Nature – the Gravitational force, the Electromagnetic force, the Weak and the Strong force – into one, giant, neat, beautiful, coherent and consistent mathematical framework that has a direct correspondence with physical and natural reality as we know it.

Despite his legendary lifelong failure to attain a unified field theory, Einstein [2, 3] understood very well the need for tensorial affine connections in the construction of a unified field theory. Einstein [2, 3] was not alone in this esoteric pot of wisdom; amongst others, towering figures of history such as Eddington [4] and Schrödinger [5–7] all but made similar noteworthy attempts to attain a unified field theory that made use of tensorial affines.

In the present work, this idea of tensorial affine connections is a fundamental lynchpin in the construction of what we believe is a noteworthy stepping stone to a *Final Unified Field Theory* (FUFT) of the gravitational phenomenon and the other forces of Nature. When we here say Final Unified Field Theory, we mean this in the context of the path (see [8–10]) that we are pursuing in order to arrive at what we believe is the FUFT.

In order for us to give the reader the correct scope of the present work, we must hasten and say that the present work is part and parcel of our upcoming monograph on this grand dream of Einstein. What we present herein is but a portion thereof. We herein demonstrate that gravitomagnetism has a fundamental geometric justification well within the scheme of Weyl's [1] supposed failure. We strongly believe – or are of the innate view – that the much sought for path to a successful *Quantum Geometrodynamic* (QGD) theory will be achieved very soon *via* a recasting of the gravitational phenomenon into a Maxwell-type formalism where the quantization of the gravitational field will prove to be the trivial exercise of quantizing a four-vector field $\mathcal{A}_\mu$ associated with the gravitational field. Through the well known quantization

procedures discovered in the quantization of the electromagnetic four-vector field in *Quantum Electrodynamics* (QED), the gravitational four-vector field can be quantized too in this very same manner.

We must say that our theory is directly inspired by Weyl's geometry [1] – a geometry that for the first time made the great and esoteric stride and endeavour to bring the electromagnetic and gravitational forces together into a fruitful and harmonious union that did not last beyond Einstein's first criticism of it (see e.g. [11]). Unlike what we have done in our previous work (in [8–10]), we shall not anymore bother our reader with the plethora of the exciting and fascinating historic anecdotes associated with the pursuit of a unified field theory that brings the gravitational and quantum phenomenon into one giant, neat, coherent and consistent mathematical framework. We deal here directly with the purest portions and jewels of our effort.

In their noble quest and search for a unified field theory of the quantum and gravitational phenomenon, physicists – and mathematicians alike – have been motivated by various reasons. In our case, our motivation has been, and is solemnly to overcome the obvious great difficulty associated with the General Theory of Relativity (GTR)'s geodesic equation of motion, namely:

$$\frac{d^2 x^\lambda}{ds^2} - \Gamma^\lambda_{\alpha\delta} \frac{dx^\alpha}{ds} \frac{dx^\delta}{ds} = 0 \qquad (1)$$

where $ds = cd\tau$ is the line element, τ is the relativistic proper time, $c = 2.99792458 \times 10^8 \, \mathrm{m\,s^{-1}}$ (CODATA 2018) is the speed of light in *vacuo*, x^μ is the four-position of the particle in spacetime, and $\Gamma^\lambda_{\mu\nu}$ are the usual Christoffel three-symbols [12]*. Because of the non-tensorial nature of the affine connection $\Gamma^\lambda_{\mu\nu}$, this geodesic (1) of motion does not hold fast – in the truest sense – to the depth of the letter and essence of the philosophy deeply espoused and embodied in Einstein's *Principle of Relativity* (PoE) [13], namely that physical laws must require no special set of coordinates where they are to be formulated.

The non-tensorial nature of the affine connection requires that the equation of motion must first be formulated in special kind of coordinate systems known as a *geodesic coordinate system*†, yet the PoE forbids this. This problem has never been adequately addressed in the GTR. As Einstein [2] noted, a permanent way out of this dilemma is to find a geometry whose affine connections are tensors. This is what we do herein. At the end of our quest – for the gravitational

---

*These symbols are named after German mathematician and physicist Elwin Bruno Christoffel (1829-1900). Christoffel first introduced these symbols in a paper on differential forms in n variables, published in Crelle's Journal: see [12].

†A geodesic coordinate system is one in which the Christoffel three-symbols $\Gamma^\lambda_{\mu\nu}$ vanish at all points on the given set of coordinates – i.e. $\Gamma^\lambda_{\mu\nu} = 0$. An example is the flat rectangular (x, y, z) system of coordinates. However, when one moves from this (x, y, z) rectangular system of coordinates to say the spherical (r, θ, φ), the resulting affine $\Gamma^{\lambda'}_{\mu'\nu'}$ is not zero – i.e. $\Gamma^{\lambda'}_{\mu'\nu'} \neq 0$.

phenomenon as a whole – we arrive not by design, but rather by serendipity, at a *gravitomagnetic theory* similar to that of Maxwell [14].

In current efforts being made on both the theoretical (in e.g. [15–19]), and observational front (in e.g. [20–24]), gravitatomagnetism is predominately understood in the context of Einstein's [25–28] linearised first order approximation of the GTR. Our approach is different to this predominant approach.

We herein consider gravitomagnetism as an exact theory independent of the GTR in much the same way it was conceived by Maxwell [14] and Heaviside [29, 30] and further championed (in modern times) e.g. by Jefimenko [31] and Behera [32] amongst others. The present gravitomagnetic theory falls within the realm of a more ambitious attempt that we are currently working on, i.e. an attempt at an all-encompassing *Unified Field Theory* (UFT) of all the forces of Nature (see [10, 33]). We shall say nothing about this attempt but direct the interested reader to these works.

In closing this introductory section, we shall give a brief synopsis of the remainder of the paper. In §2, we give a brief exposition of the GTR. In §4, we give an exposition of Weyl's theory [1]. In §3, we give a non-geometric justification of gravitomagnetism. In §5, we present our theory. Thereafter, in §6, in preparation for the presentation of the gravitomagnetic field equations, we express the new affine ($\bar{\Gamma}^\lambda_{\mu\nu}$) and the Riemann tensor ($R_{\mu\nu}$) in terms of the gravitational Maxwell-type field tensor ($\mathfrak{F}_{\mu\nu}$). Therein §6, we also work out the geometrically derived material tensor ($\mathcal{T}_{\mu\nu}$) so that its terms correspond with what we know from the physical world. In §7, we write down the resultant field equations. Lastly, in §8, a general discussion is given.

2 Brief exposition of the GTR

As is well known, Einstein's *Special Theory of Relativity* [34] deals with inertial observers while the GTR deals with non-inertial observers. The problem with non-inertial observers is that gravitation becomes a problem since it is an all pervading *non-vanishing force*. By analysing the motion of a test body in free fall motion in a gravitational field, Einstein [13] was able to overcome this problem of gravitation by noting that if the gravitational (m_g) and inertia mass (m_i) were equal or equivalent, then gravitation and acceleration are equivalent too. This meant that the effect(s) of acceleration and gravitation are the same. One can introduce or get rid of the gravitational field by introducing acceleration into the system. Because of the importance of this, it came to be known as the *Principle of Equivalence*, and Einstein [25] took this as a foundational pillar to be used for the construction of his GTR.

2.1 Principle of Equivalence

The deep rooted meaning of the Principle of Equivalence is that physical laws should remain the same in a local reference system in the presence of a gravitational field as they do in

an inertial reference system in the absence of gravitation. In Einstein's own words [13]:

> **Einstein's Principle of Equivalence (PoE):** We shall therefore assume the complete physical equivalence of a gravitational field and the corresponding acceleration of the reference system. This assumption extends the Principle of Relativity to the case of uniformly accelerated motion of the reference system.

A consequence of this is that no mechanical or optical experiment can locally distinguish between a uniform gravitational field and uniform acceleration. It is here that we would like to point out that the PoE as used in the formulation of the GTR does not demand that the physics must remain invariant. By "the physics" we mean that the description of a physical event ought to remain invariant unlike, for example, in black hole physics – where, depending on the coordinate system employed (and not the reference system – this is important), a particle can be seen to pass or not pass through the Schwarzschild sphere for the same observer supposedly under the same conditions of experience. Also the chronological ordering of events is violated – i.e. the *Law of Causality* is not upheld.

For example, as first pointed out by the great mathematician, logician and philosopher Kant Gödel [35], in a rotating Universe, it is possible to travel back in time, invariably meaning to say it is possible in principle to violate the *Second Law of Thermodynamics*. Though the idea of time travel is very fascinating and appealing to the mind, it is difficult to visualize by means of binary logical reasoning how it can work in the physical world as we know it. From intuition, the laws of Nature must somehow have deeply engraved and embedded in them the non-permissibility of time travel. We believe that such illogical outcomes emerging out from a legitimate application of the laws of Nature can be solved if the geometry of the Universe is built on tensorial affinities.

2.2 Generalized Principle of Equivalence

Therefore, in order to avoid physical absurdities emerging from supposedly well-founded laws of Nature, we must demand of our theories that "the physics" emerging from the theory, that is to say, the physical state and the chronological ordering of events, must remain invariant – i.e. we must extend the Principle of Equivalence to include the physical state or physical description of events and as well the Law of Causality. Because this must be universal and important, let us call the extended Principle of Equivalence, the *Generalized Principle of Relativity*:

> **Generalized Principle of Relativity (GPR):** Physical laws have the same form in all equivalent reference systems independently of the coordinate system used to express them and the complete physical state or physical description of an event emerging from these

laws in the respective reference systems must remain absolutely and independently unaltered – i.e. invariant and congruent – by the transition to a new coordinate system.

This forms the basic guiding principle of the present theory. The deeper meaning of the GPR is that, if one is describing the same physical event in spacetime e.g. a black hole, it should not be permissible to transform away a singularity by employing a different set of coordinates as is common place in the study of the Schwarzschild metric. If the singularity exists, it exists independently of the coordinate system and reference system used – it is intrinsic and permanent, it must exist at all levels of the theory.

Therefore, if we are to have no singularities, the theory itself must be free of these. If a particle is seen not to pass through the event horizon, it will not be seen to pass through the boundary of the event horizon no matter the coordinate system employed and the reference system to which the current situation is transformed into. In order for this, there is need for the affine connections to be tensors and this is what we shall try to achieve in the present – i.e. a geometry endowed with tensorial affine connections. For completeness, self-containment and latter instructive purposes, in the next subsection, we will take a look at the non-tensor affine connections of Riemann geometry.

2.3 Affine connection

Now, back to the main vein: the Principle of Equivalence is, in the context of Riemann geometry, mathematically embodied in the equation:

$$g_{\mu\nu;\alpha} = g_{\mu\nu,\alpha} - \Gamma^{\delta}_{\alpha\mu}g_{\delta\nu} - \Gamma^{\delta}_{\alpha\nu}g_{\mu\delta} = 0 \qquad (2)$$

where $g_{\mu\nu}$ is the metric tensor describing the geometry of spacetime and $\Gamma^{\lambda}_{\mu\nu}$ is the affine connection. This affine connection is obtained as follows (e.g. [36, pp. 59–60]): first we write down two equations obtained by way of right-cyclically permuting the $\mu\nu\sigma$-indices in (2) for the term $g_{\mu\nu,\sigma}$, i.e.:

$$g_{\mu\nu;\alpha} = g_{\mu\nu,\alpha} - \Gamma^{\delta}_{\alpha\mu}g_{\delta\nu} - \Gamma^{\delta}_{\alpha\nu}g_{\mu\delta} = 0, \qquad (3)$$

$$g_{\mu\nu;\alpha} = g_{\mu\nu,\alpha} - \Gamma^{\delta}_{\alpha\mu}g_{\delta\nu} - \Gamma^{\delta}_{\alpha\nu}g_{\mu\delta} = 0. \qquad (4)$$

Second, we now subtract from (2) the sum of (3) and (4), and use the symmetry of the connection ($\Gamma^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\nu\mu}$) and as well of the metric ($g_{\mu\nu} = g_{\nu\mu}$) to obtain: $\left(g_{\mu\nu,\alpha} - g_{\alpha\mu,\nu} - g_{\nu\alpha,\mu}\right) + 2g_{\alpha\delta}\Gamma^{\delta}_{\mu\nu} = 0$, hence:

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}g^{\delta\lambda}\left(g_{\delta\mu,\nu} + g_{\nu\delta,\mu} - g_{\mu\nu,\delta}\right). \qquad (5)$$

The affine connections play an important role in that they relate tensors between different reference systems and coordinate systems. Its drawback insofar as physical laws are con-

cerned is that it is not a tensor. It transforms as follows:

$$\Gamma^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^{\delta}} \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} \Gamma^{\delta}_{\mu\nu} + \underbrace{\frac{\partial x^{\lambda'}}{\partial x^{\delta}} \frac{\partial^2 x^{\delta}}{\partial x^{\mu'} \partial x^{\nu'}}}_{\textbf{extra term}} \; . \tag{6}$$

The extra term on the right makes it a non-tensor and without it, the Christoffel symbol would be a tensor. Most of the problems facing the GTR can be traced back to the non-tensorial nature of the affine connections. Some of the problems will be highlighted in the succeeding section. Due to the nature of these affinities, the real problem is that in its bare form, Riemann geometry does not provide a way to determine permissible and non-permissible coordinate and reference system transformations. The new hybrid geometry on which the UFT being championed is built, does have a way to determine permissible and non-permissible coordinate and reference system transformations.

2.4 Line element

Now, both the invariance and covariance of physical laws under a change of the coordinate system and/or reference system transformation is, in Riemann geometry, encoded and/or expressed through the invariance of the line element: $ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu}$. The line element is a measure of the distance between points in spacetime and remains invariant under any kind of transformation of the reference system and/or the coordinate system. This is the essence of the GTR. From this, Einstein was able to deduce that gravitation is and/or can be described by the metric tensor $g_{\mu\nu}$ thus, according to the Einstein doctrine of gravitation, it (gravitation) manifests itself as the curvature of spacetime. Through his (Einstein) own intuition and imagination, he was able to deduce that the curvature of spacetime ought to be proportional to the amount of matter-energy present in spacetime — a fact that has since been verified by numerous experiments.

2.5 Einstein's field equations

The resulting gravitational law emerging from Einstein's thesis stated above – namely that the curvature of spacetime should be proportional to the amount of matter-energy present in spacetime – is:

$$\overbrace{\underbrace{R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu}}_{\textit{beautiful and splendour}}}^{\textbf{MARBLE}} = \overbrace{\underbrace{\kappa_E T_{\mu\nu}}_{\textit{ugly and loathsome}}}^{\textbf{WOOD}} \tag{7}$$

where $\kappa_E = 8\pi G/c^4$ is Einstein's constant of gravitation, $G = 6.67430(15) \times 10^{-11} \, \text{kg}^{-1} \, \text{m}^3 \, \text{s}^{-2}$ (CODATA 2018) is Newton's universal constant of gravitation, $R_{\mu\nu}$ is the contracted Riemann curvature tensor, R is the Ricci scalar, and $T_{\mu\nu} = \varrho_g v_{\mu} v_{\nu} + p g_{\mu\nu}$ is the stress and energy tensor where ϱ_g is the density of matter, p is the pressure, v_{μ} the four-velocity, and Λ

is the controversial ad hoc *Cosmological Constant* term added by Einstein [37] so as to stop the Universe from expanding. Einstein [37] was motivated to include the cosmological constant because of the strong influence from the astronomical wisdom of his day that the Universe appeared to be static and thus was assumed to be so.

In the later years of his scientific life while in hot pursuit of a unified field theory – according to his official scientific biographer – Abraham Pais [38], Einstein would look at his equation (7) and compare the left-hand side with marble and the right-hand side with wood, and he would admire the marble side calling it *beautiful and splendour* and, on looking at the right-hand side, he would be filled with sadness whereby he would moan calling it *ugly and loathsome*. His prime and hence immediate goal therefore (see e.g. [39]) was to turn the *ugly wood* into *beautiful marble*.

All Einstein hoped for and wanted in his quest, was that all the fields including the material field $T_{\mu\nu}$, be derived from pure geometry, rather than "glue" the two seemingly independent parts (i.e. the curvature $R_{\mu\nu} - Rg_{\mu\nu}/2$ and material tensor $T_{\mu\nu}$) *via* some mere constant κ_E. Einstein was extremely dissatisfied with this state of affairs [38] and thus hoped that a theory would be found in the future where the material tensor is derived directly from the geometry as a direct consequence of the geometry itself. We must say, that, if our ideas prove themselves worthy, it appears we have just managed to derive the material fields from the *Resultant World Geometry*.

3 Present justification of gravitomagnetism

For example, take *Maxwell's Five Equations of Electrodynamics* [14] – i.e. the typical four equations that we are used to involving the reciprocal E and B-fields plus the *Law of Conversation of Electric Charge and Current*. Certainly, to a foremost theoretical physicist such as Paul Dirac (see e.g. [40–42]), these equations are without doubt beautiful in every aspect of the word beauty; and to seal the matter, their foundations are well verified and anchored in experience. But asking what is the fundamental basis for their existence led José Heras [43] to the tangibly solid mathematical fact that Maxwell's equations [14] are nothing more than a consequence of the conservation of electronic charge. That is to say, what you need for the existence of Maxwell's equations [14] is just the conservation of electric charge and current; nothing more and nothing less. Surely – to say that only the conservation of electronic charge and current is all that is needed for Maxwell's Equations to exist – this is certainly deep, isn't it?

Given that the gravitational mass – which is responsible for gravitation – follows a similar law of conversation in the form of the conservation of mass-energy and momentum, rather trivially, one can easily extend this to the gravitational phenomenon and justify the need for gravitomagnetism. Heras [43] did not make this trivial and obvious extrapolation. In addition to this, we must say that we have not

seen in the most recent literature any attempt to use Heras' [43] existence theorem to justify gravitomagnetism. However, by way of analogy with the equations of electrodynamics given the similarity between Newton and Coulomb's inverse square laws, Maxwell [14] and Heaviside [29, 30] already had introduced gravitomagnetism. Sadly, because of lack of experimental backing, gravitomagnetism derived in this way has largely been treated as an endeavour belonging to the realm of *pseudo-science*, rather than science. Many scientists that have followed in an effort to try and investigate this gravitomagnetic phenomenon have struggled to shrug-off the pseudo-science tag hanging at the nimbus of gravitomagnetism.

In the present section, we are going to give a brief exposition of Heras [43] and Behera's [32] existence theorems. These theorems are enough to convince sceptics that like electricity and magnetism, the gravitational phenomenon aught to be described by a four-vector potential. In addition to Heras [43] and Behera's [32] existence theorems, this paper will add a purely geometric justification and this geometric justification follows the same geometric path as the GTR wherein the gravitational phenomenon is described by the metric. Because this demonstration – that we are going to give of the geometric justification of gravitomagnetism – uses the modern description of gravitation as a metric phenomenon, it certainly is not far-off in its outlook, vision and conception with the modern idea of a metric description of gravity. Surely, this aspect of the present ideas must – somehow – make the ideas propagated herein appeal to the reader. In the next subsection, we shall give an exposition of Heras' theorem [43].

3.1 Heras's (2007) existence theorem

In a nutshell, Heras [43] formulated – what in our view is – a very important *Existence Theorem* that states that, given any space and time-dependent localized scalar and vector sources satisfying the continuity equation – as is the case with electromagnetism – there exists in general, two retarded vector fields (X, Y) that satisfy a set of four field equations that are similar in nature and form to Maxwell's equations. By applying the theorem to the usual electrical charge and current densities, the two retarded fields are identified with the reciprocal electric (E) and magnetic (B) fields and the associated field equations with Maxwell's equations [14], i.e.: $X := E, Y := B$.

In brief, what Heras [43] proved is that if ϱ_e is the charge density and $\vec{J}$ is the associated current corresponding to this charge, i.e.:

$$\frac{\partial \varrho_e}{\partial t} = -\vec{\nabla} \cdot \vec{J}, \tag{8}$$

then, there must exist two corresponding fields, X and Y, that

satisfy the following set of equations:

$$
\begin{aligned}
\vec{\nabla} \cdot X &= \alpha \varrho_e & \text{(a)} \\
\vec{\nabla} \cdot Y &= 0 & \text{(b)} \\
\vec{\nabla} \times X + \gamma \frac{\partial Y}{\partial t} &= 0 & \text{(c)} \\
\vec{\nabla} \times Y - \frac{\beta}{\alpha} \frac{\partial X}{\partial t} &= \beta \vec{J} & \text{(d)}
\end{aligned}
\tag{9}
$$

where α, β, γ are arbitrary positive constants and are related to the speed of light c by the equation $\alpha = \beta \gamma c^2$. In the case of electricity and magnetism, if X and Y are identified with the electric and magnetic fields respectively, then we will have Maxwell's classical equations [14] for electrodynamics – in which case $\alpha = 1/\varepsilon$, $\beta = \mu$, and, $\gamma = 1$. Clearly, this axiomatic and fundamental approach of deriving Maxwell's field equations [14] strongly suggests that electric charge and current conservation – and nothing else – can be considered to be the most fundamental assumption underlying Maxwell's equations [14] of electrodynamics. Next, we give an exposition of Behera's [32] theorem.

3.2 Behera's (2006) theorem

Using the Law of Conservation 0f Mass-Energy-Momentum and the Poisson-Laplace equation (10), the endeavour of the present section is to demonstrate – as Behera [32] did – that much the same as the Coulomb electrical potential, the Newtonian gravitational potential ϱ_g has an associated vector field. We shall denote this vector field by the symbol A_g and we shall call it the gravitomagnetic vector potential and in short we shall call it the *g-magnetic* vector potential. This fact that we can associate ϱ_g with A_g has been known for a considerable amount of time now. That is, for more than a century (≥ 120 years), it has been known (since Heaviside [29, 30]) that the inclusion of a magnetic-like vector field in Newtonian gravitational theory can be justified from two immutable facts (see e.g. Behera [32]), i.e. from the Poisson-Laplace equation for gravitation, namely:

$$\vec{\nabla} \cdot \vec{g} = -4\pi G \varrho_g \tag{10}$$

where $\vec{g}$ is the gravitational field intensity at a given point in the gravitational field, ϱ_g is the gravitational potential, and from the equation of conservation of mass-energy and momentum, namely: $\partial \varrho_g / \partial t = -\vec{\nabla} \cdot \vec{\Im}$, where $\vec{\Im} = \varrho_g v$, is the momentum density with v being the velocity of the material whose density is ϱ_g.

In order to see this, from (10) we know very well that: $\dot{\varrho}_g = -(1/4\pi G)(\vec{\nabla} \cdot \dot{\vec{g}})$. Let us set: $\widetilde{\mu} = 1/4\pi G$, so that: $\dot{\varrho}_g = (1/4\pi G)(\vec{\nabla} \cdot \dot{\vec{g}})$ can now be written as: $\dot{\varrho}_g = -\widetilde{\mu} \vec{\nabla} \cdot \dot{\vec{g}}$. From this, it follows that:

$$\frac{\partial \varrho_g}{\partial t} = -\vec{\nabla} \cdot \vec{\Im} = -\vec{\nabla} \cdot \left(\widetilde{\varepsilon} \frac{\partial \vec{g}}{\partial t} \right), \tag{11}$$

hence:

$$\vec{\nabla} \cdot \left[\widetilde{\varepsilon} \frac{\partial \vec{\mathfrak{g}}}{\partial t} + \vec{\mathfrak{J}} \right] = 0. \tag{12}$$

Now, it is a *bona fide* mathematical fact that for any general vector say $\vec{\mathfrak{B}} = \vec{\mathfrak{B}}(x)$, the following holds always:

$$\vec{\nabla} \cdot \left(\frac{\vec{\nabla} \times \vec{\mathfrak{B}}}{\widetilde{\mu}} \right) \equiv 0. \tag{13}$$

where $\widetilde{\mu}$ is a constant – which, akin to the electromagnetic permeability (μ_0) and permittivity (ϵ_0) of free space, we shall define this constant $\widetilde{\mu}$ is such that: $\widetilde{c} = 1/\sqrt{\widetilde{\mu}\widetilde{\varepsilon}}$, where $\widetilde{c}$ is the speed of gravity in free space. By comparing (12) and (13), it follows that:

$$\frac{\vec{\nabla} \times \vec{\mathfrak{B}}}{\widetilde{\mu}} = -\vec{\mathfrak{J}} + \widetilde{\varepsilon} \frac{\partial \vec{\mathfrak{g}}}{\partial t}. \tag{14}$$

What this really means is that the gravitational field $\vec{\mathfrak{g}}$ has an associated magnetic-like field $\vec{\mathfrak{B}}$. Hence, one can make the very bold conclusion that the very laws of Nature (10) and $\partial \varrho_g / \partial t = -\vec{\nabla} \cdot \vec{\mathfrak{J}}$ invariably imply an associated magnetic-like field for the gravitational field. Following tradition, we shall call this magnetic-like field the gravitomagnetic field and for short, we shall call it the *g-magnetic* field.

Now, (10) and (14) have a seductive and irresistible resemblance with Maxwell's source-coupled equations so much so that for the brave that have set their mind on this, they have proceeded without detouring to make a complete formal analogue with Maxwell's equations [14], in which process, the phenomenon known as *gravitomagnetism* found its original birth certificate. Therefore, as a complete set, the *Field Equations of Gravitomagnetism*, are:

$$\vec{\nabla} \cdot \vec{\mathfrak{g}} = -\varrho_{\mathrm{g}}/\widetilde{\varepsilon} \qquad \text{(a)}$$

$$\vec{\nabla} \times \vec{\mathfrak{g}} = -\frac{1}{\widetilde{c}} \frac{\partial \vec{\mathfrak{B}}}{\partial t} \qquad \text{(b)}$$

$$\vec{\nabla} \cdot \vec{\mathfrak{B}} = 0 \qquad \text{(c)} \qquad (15)$$

$$\vec{\nabla} \times \vec{\mathfrak{B}} = -\widetilde{\mu}\vec{\mathfrak{J}} + \frac{1}{\widetilde{c}^2} \frac{\partial \vec{\mathfrak{g}}}{\partial t}. \qquad \text{(d)}$$

This completes our exposition of the non-geometric justification of gravitomagnetism. In the next section, we shall for self-containment and latter instructive purposes, present a brief exposition of Weyl's theory [1] and in the penultimate thereof, we present our partial modification of it.

4 Weyl geometry

In §4.1, we give a brief exposition of Weyl's geometry [1] and thereafter in §4.2, we present the *New Weyl Geometry* (NWG) upon which the proposed gravitomagnetic theory is based.

4.1 Original Weyl geometry

By way of addition of a conformal factor $e^{2\phi}$ to the metric $g_{\mu\nu} \mapsto e^{2\phi} g_{\mu\nu}$, Weyl [1] built his geometry by supplementing the Christoffel affine connection $\Gamma^\lambda_{\mu\nu}$ of Riemann geometry with a tensorial affine $\mathcal{W}^\lambda_{\mu\nu}$:

$$\mathcal{W}^\lambda_{\mu\nu} = g^\lambda_{\ \mu} \mathcal{A}_\nu + g^\lambda_{\ \nu} \mathcal{A}_\mu - g_{\mu\nu} \mathcal{A}^\lambda, \tag{16}$$

where $\mathcal{A}_\mu$ is a four-vector that Weyl [1] had to define as the electromagnetic four-vector appearing in Maxwell's theory of electrodynamics [14].

In Weyl's geometry [1] where the length of vector changes from point to the next (see e.g. [33]), the new affine connection $\overline{\Gamma}^\lambda_{\mu\nu}$ (or Christoffel-Weyl connection) is given by:

$$\overline{\Gamma}^\lambda_{\mu\nu} = \Gamma^\lambda_{\mu\nu} + \mathcal{W}^\lambda_{\mu\nu}. \tag{17}$$

The transformational properties of the new Christoffel-Weyl affine connection $\overline{\Gamma}^\lambda_{\mu\nu}$ are identical to those of the original Christoffel three-symbol $\Gamma^\lambda_{\mu\nu}$. So, from a "transformational properties" (topological) standpoint, Weyl's [1] addition is justified.

The versatile and agile Weyl [1] was quick to note that this new Christoffel-Weyl affine (17) is invariant under the following rescaling of the metric $g_{\mu\nu}$ and the four-vector $\mathcal{A}_\mu$:

$$
\begin{aligned}
g_{\mu\nu} &\longmapsto e^{2\chi} g_{\mu\nu} \\
\mathcal{A}_\mu &\longmapsto \mathcal{A}_\mu + \kappa_0^{-1} \partial_\mu \chi
\end{aligned}
, \tag{18}
$$

where $\chi = \chi(\boldsymbol{r}, t)$ is a well behaved, arbitrary, smooth, differentiable, integrable and uniform continuous scalar function, and κ_0 is a constant with the dimensions of inverse length. This constant κ_0 has been introduced for the purposes of dimensional consistency, since we here assume that the four-vector $\mathcal{A}_\mu$ and the true scalar χ are dimensionless physical quantities.

Now, because Maxwell's electromagnetic theory [14] is invariant under the same gauge transformation which the four-vector $\mathcal{A}_\mu$ has been subjected to in (18), the great mind of Weyl seized the golden moment and identified this four-vector $\mathcal{A}_\mu$ with the electromagnetic four-vector potential. Weyl went on to assume that the resulting theory was a unified field theory of gravitation and Maxwellian electrodynamics. Weyl's hopes where dashed – first, starting with Einstein's lethal critique of the theory. Later, others joined Einstein in their merciless critique and dismissal of Weyl's theory [1], where they argued that despite its irresistible grandeur and exquisite beauty, Weyl's theory [1] cannot possibly describe the measured reality of our present world.

4.2 New Weyl geometry

Despite the many ingenious attempts (starting with e.g. Weyl [44, 45]) to rework and revive it over the course of time since
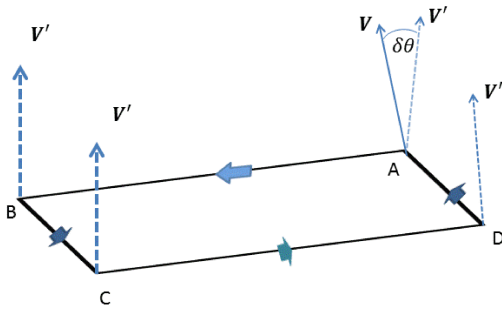
Fig. 1: **Parallel Transport:** The vector V is parallel transported in a closed circuit. Upon arrival at its original position, the vector is not equal to the original vector and this is a result of the curvature of the space in question.

its inception, and its apparent refusal to go away as evidenced by the continued interest* in this beautiful geometry of Weyl [1], it is a generally accepted view that as a basis for a physical theory, Weyl's [1] arcanely beautiful geometry exists beyond redemption. This geometry is the geometry on which [1] made his attempt – the first such – on a UFT of the gravitational and electromagnetic fields. Against this deeply entrenched belief in the non-redeemability of the Weyl [1] geometry into something with a bearing and correspondence with physical and natural reality, we made in [33] the endeavour of calling forth this theory out of the tomb where it was resting. In the present, we go further to give it a perdurable fresh breath of life.

As pointed out by e.g. Schrödinger and Einstein [3, 5–7] and is well known, is that – tensorial affine connections preserve both the length and direction of a vector upon parallel transport. The Christoffel symbols of Riemann geometry preserve only the length and the angle changes from one point the next and this is where the issue with Einstein's GTR [55] lies. Preservation of both the length and angle of a vector upon parallel transport has always been known to be a fundamental key to the attainment of a truly generalized Theory of Relativity [56, 57].

The proposed RWS is a spacetime which preserves both the length and direction of a vector upon parallel transport. As shown in Fig. 1, say the vector V is transported in a closed circuit such that it returns to its original position and V' is the resulting vector after parallel transport; in normal Riemann geometry, while $|V| = |V'|$, the angle $\delta\theta$ between these two vectors, while it can in some cases equal zero, is not necessarily zero i.e. $V \cdot V' \neq 0$. However, on the RWS, we have for all points of space and time on this spacetime the constraints $|V| = |V'|$ and $V \cdot V' = 0$: i.e. both the length and direction of a vector are preserved upon parallel transport of any vector.

_____

*See e.g. [46–54].

The preservation of both the length and angle on the RWS is attained by requiring that the affine connections of this spacetime be tensors. As far as we can tell from our wide ranging searches across the length, breath and depth of the available literature on unified theories (cf. [56, 57]), with the failure to obtain tangible results on this front, the idea of tensorial affinities as key to the attainment of a unified field theory seems to have naturally fallen on the wayside with very few – if any – researchers taking it up. As one will be able to judge for themselves and by themselves, the novelty of our approach lies in our treatment of the unit vectors.

As pointed out in the instance of (18), to attain the desired tensorial affinities, we noted that Weyl [1] had built his very beautiful but failed unified field theory of gravitation and electromagnetism on a pseudo-Riemann spacetime that is invariant under the re-gauging of the metric from $g_{\mu\nu}$ to $e^{2\chi}g_{\mu\nu}$: i.e. after the transformation $g_{\mu\nu} \longmapsto e^{2\chi}g_{\mu\nu}$, the field equations of the resulting geometry or theory thereof remain unaltered provided the four-vector of his theory $\mathcal{A}_\mu$ also underwent the following gauge transformation: $\mathcal{A}_\mu \longmapsto \mathcal{A}_\mu + \kappa_0^{-1}\partial_\mu\chi$. The mathematical structure of the resulting Weyl unified field theory, insofar as the properties of the affine connections is concerned, this theory – despite its elegant introduction of a four-vector field – has the same topological deformations as the original Riemann spacetime.

4.2.1 Riemann-Weyl metric

As already pointed out in §4.1, Weyl added a tensor $\mathcal{W}_{\mu\nu}^\lambda$ to the Christoffel three-symbol $\Gamma_{\mu\nu}^\lambda$, that is to say, if $\widetilde{\Gamma}_{\mu\nu}^\lambda$ is the new Christoffel symbol for the Weyl space, then:

$$\widetilde{\Gamma}_{\mu\nu}^\lambda = \Gamma_{\mu\nu}^\lambda + \mathcal{W}_{\mu\nu}^\lambda. \tag{19}$$

Because $\mathcal{W}_{\mu\nu}^\lambda$ is a tensor, the fundamental transformational properties of the new Christoffel three-symbol $\widetilde{\Gamma}_{\mu\nu}^\lambda$ are the same as the old Christoffel three-symbol $\Gamma_{\mu\nu}^\lambda$; therefore, the Weyl space inherits the same topological and structural defects and problems of the Riemann spacetime – that is, problems to do with non-tensorial affinities.

In [33], for the metric of the RWS $\overline{g}_{\mu\nu}$, instead of making it conformal only at the instance of a gauge transformation, we chose that it ($\overline{g}_{\mu\nu}$) be intrinsically and inherently conformal. That is to say, the fundamental metric $\overline{g}_{\mu\nu}$ of the RWS be such that $\overline{g}_{\mu\nu} = \varrho g_{\mu\nu}$, where $g_{\mu\nu}$ remains as the metric of the usual Riemann spacetime and this metric is what is used on the RWS to raise and lower the Greek indices ($\mu\nu\ldots$) just as happens in normal Riemann spacetime. In Weyl's theory [1], the function ϱ is a scalar. However in [33], this function takes a decisive new role: …it (the scalar χ) must – for better or for worse, yield in the favour of our desideratum – i.e. it must yield for us nothing but tensorial affinities. This is our quest, desire and uncompromising demand.

Thus, in recasting Weyl's theory [1] so that it overcomes once and for all-time Einstein's criticism, we will not take the traditional route that was taken by Weyl [1] because in so doing, we will fall into the same trap which the great Weyl fell victim to. At our point of departure, we wave goodbye to Riemann geometry and efferently prepare to embrace a totally new geometry, a hybrid Riemann geometry which has the same feature as Weyl [1], less of course the change of length of vectors under transformations or translations. The route that we are about to take is equivalent and the reason for changing the sails is that the present route allows us to demonstrate later how Weyl would have overcome Einstein's critique that gave the theory a still birth. Actually, this route allows us to pin down exactly where Weyl's theory [1] makes an unphysical assumption.

4.2.2 Pseudo-scalar and affine vector

In mathematics – linear algebra in particular – a pseudo-scalar is a function which upon a transformation of the coordinate system behaves like a true scalar – albeit – upon a parity transformation, it changes sign (see e.g. [58, 59]). A true scalar does not do this, it remains invariant. As has already been made clear in the exposition of Weyl's theory [1] is the fact that one of the most powerful ideas in physics is that physical laws do not change when one changes the coordinate system used to describe these physical laws. The fact that a pseudo-scalar reverses its sign when the coordinate axes are inverted clearly suggests that these objects are not the best objects to describe a physical quantity, as this could percolate to the physical laws themselves.

Now, because we want to introduce a new kind of pseudo-scalar that will help us in our endeavours to obtain tensorial affinities, in order to distinguish this new and soon to be defined pseudo-scalar from the above described pseudo-scalar, we shall call the above described pseudo-scalar a *pseudo-scalar of the first kind*, and the new pseudo-scalar to be defined shortly, a *pseudo-scalar of the second kind*. To that end, we shall hereafter start off by defining a "new" mathematical object, $\mathcal{V}_\mu$, that we shall call an *affine vector*. This quantity, $\mathcal{V}_\mu$, is the derivative of the dot-product of an arbitrary four-vector $\mathcal{B}_\lambda$ and the (non-arbitrary) four-position x^λ i.e.:

$$\mathcal{V}_\mu = \frac{\partial_\mu \mathcal{S}}{\mathcal{S}} = \partial_\mu \ln \mathcal{S} \qquad (20)$$

where:

$$\mathcal{S} = \mathcal{B}_\delta x^\delta. \qquad (21)$$

From (20) and (21), it follows that:

$$\mathcal{V}_\mu = \mathcal{B}_\mu + \frac{x^\delta \partial_\mu \mathcal{B}_\delta}{\mathcal{S}}. \qquad (22)$$

Clearly, upon a coordinate and/or transformation of the reference system, the vector-like quantity $\mathcal{V}_{\mu'} = \partial_{\mu'} \mathcal{S}'/\mathcal{S}'$ is related

to $\mathcal{V}_\mu$ as follows:

$$\mathcal{V}_{\mu'} = \frac{\partial x^\delta}{\partial x^{\mu'}} \mathcal{V}_\delta + \frac{\partial^2 x^\delta}{\partial x^{\mu'} \partial x^{\Omega'}}. \qquad (23)$$

From (23), we see that the quantity $\mathcal{V}_\mu$ transforms like the affine tensor hence our calling it – affine vector. The scalar $\mathcal{S}$ in (21) is what we shall define as a *pseudo-scalar of the second kind*. Such a scalar has the property that its four-position derivative is not a four-vector as one would expect in the case of a true scalar. In the next section, we shall now consider the Riemann-Weyl covariant derivative in the light of the new mathematical object that we have just defined – i.e. the pseudo-scalar of the second kind.

4.2.3 Riemann-Weyl covariant derivative

Taking into account the above defined pseudo-scalar of the second kind, as we consider the Riemann-Weyl covariant derivative, we will begin with the usual Riemann covariant derivative $g_{\mu\nu;\sigma} = 0$ of Riemann geometry. As already alluded, the condition $g_{\mu\nu;\sigma} = 0$ is the foundation stone of Riemann geometry. We will uphold this covariant derivative condition under the Weyl conformal transformation $g_{\mu\nu} \longmapsto \overline{g}_{\mu\nu} = \varrho g_{\mu\nu}$ of the metric i.e. $\overline{g}_{\mu\nu;\sigma} = 0$. Likewise, the condition $\overline{g}_{\mu\nu;\sigma} = 0$ is to be taken as the foundation stone of the new hybrid Riemann-Weyl geometry. Written in full, the equation $\overline{g}_{\mu\nu;\sigma} = 0$ is given by:

$$\overline{g}_{\mu\nu;\sigma} = \varrho \left[g_{\mu\nu,\sigma} + g_{\mu\nu} \left(\frac{\partial_\sigma \varrho}{\varrho} \right) - g_{\mu\delta} \overline{\Gamma}^\delta_{\nu\sigma} - g_{\delta\nu} \overline{\Gamma}^\delta_{\mu\sigma} \right] = 0 \quad (24)$$

where the "bar" on $\overline{\Gamma}^\lambda_{\mu\nu}$ has been put so that it is made clear that this affine is neither the Christoffel symbol nor the usual Weyl connection, but is the new hybrid Riemann-Weyl connection. In conformity with the definition of a pseudo-scalar of the second kind given in (21), we shall at this point set or define the ϱ-quantity as:

$$\varrho = -2\mathcal{J}_\delta x^\delta \qquad (25)$$

where $\mathcal{J}_\sigma$ is the (gravitational) four-current density. With this definition for ϱ, it follows that (24) will reduce to:

$$g_{\mu\nu,\sigma} - g_{\mu\delta} \overline{\Gamma}^\delta_{\nu\sigma} - g_{\delta\nu} \overline{\Gamma}^\delta_{\mu\sigma} = 2 \left(\mathcal{J}_\sigma + Q_\sigma \right) g_{\mu\nu} \qquad (26)$$

where $Q_\sigma = x^\delta \partial_\sigma \mathcal{J}_\delta / \varrho = x^\delta \partial_\sigma \mathcal{J}_\delta / \mathcal{J}_\delta x^\delta$. As is the case with Weyl's original geometry [1], the covariant derivative $g_{\mu\nu;\sigma}$ does not vanish since $g_{\mu\nu;\sigma} \neq 0$.

4.2.4 Calculation of the Riemann-Weyl affine

Now – we have to calculate the resulting affine connections and for this, we have to write down the three expressions that

result from an anti-clockwise cyclic permutation of the indices μ, ν and σ in $g_{\mu\nu,\sigma}$, i.e.:

$$
\begin{aligned}
g_{\mu\nu,\sigma} - g_{\mu\delta}\overline{\Gamma}^{\delta}_{\nu\sigma} - g_{\delta\nu}\overline{\Gamma}^{\delta}_{\mu\sigma} &= 2\left(\mathcal{I}_{\sigma} + Q_{\sigma}\right)g_{\mu\nu} \quad \text{(a)}\\
g_{\sigma\mu,\nu} - g_{\sigma\delta}\overline{\Gamma}^{\delta}_{\mu\nu} - g_{\delta\nu}\overline{\Gamma}^{\delta}_{\sigma\mu} &= 2\left(\mathcal{I}_{\nu} + Q_{\nu}\right)g_{\sigma\mu} \quad \text{(b)} \quad (27)\\
g_{\nu\sigma,\mu} - g_{\nu\delta}\overline{\Gamma}^{\delta}_{\sigma\mu} - g_{\delta\mu}\overline{\Gamma}^{\delta}_{\nu\sigma} &= 2\left(\mathcal{I}_{\mu} + Q_{\mu}\right)g_{\nu\sigma} \quad \text{(c)}
\end{aligned}
$$

As usual, subtracting from (27) (a) the sum of (27) (b) and (c), and making use of the symmetries of $g_{\mu\nu}$ and $\overline{\Gamma}^{\lambda}_{\mu\nu}$ (i.e. $g_{\mu\nu} = g_{\nu\mu}$ and $\overline{\Gamma}^{\lambda}_{\mu\nu} = \overline{\Gamma}^{\lambda}_{\nu\mu}$), one obtains:

$$
\begin{aligned}
g_{\mu\nu,\sigma} - g_{\sigma\mu,\nu} - g_{\nu\sigma,\mu} + g_{\sigma\delta}\left[\overline{\Gamma}^{\delta}_{\mu\nu} + g_{\delta\nu}\overline{\Gamma}^{\delta}_{\sigma\mu}\right] = \\
= 2\left[\left(\mathcal{I}_{\sigma} + Q_{\sigma}\right)g_{\mu\nu} - \left(\mathcal{I}_{\nu} + Q_{\nu}\right)g_{\sigma\mu} - \left(\mathcal{I}_{\mu} + Q_{\mu}\right)g_{\nu\sigma}\right].
\end{aligned} \quad (28)
$$

Contracting the σ-index in (28) by multiplying (28) throughout by $g^{\sigma\lambda}$ and thereafter resetting this σ-index to δ, we obtain:

$$
\begin{aligned}
-g^{\delta\lambda}\left[g_{\delta\mu,\nu} + g_{\nu\delta,\mu} - g_{\mu\nu,\delta}\right] + 2\overline{\Gamma}^{\lambda}_{\mu\nu} = \\
-2g^{\delta\lambda}\left[\left(\mathcal{I}_{\nu} + Q_{\nu}\right)g_{\delta\mu} + \left(\mathcal{I}_{\mu} + Q_{\mu}\right)g_{\nu\delta} - \left(\mathcal{I}_{\delta} + Q_{\delta}\right)g_{\mu\nu}\right],
\end{aligned} \quad (29)
$$

hence:

$$
\overline{\Gamma}^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\mu\nu} - \mathcal{W}^{\lambda}_{\mu\nu} - Q^{\lambda}_{\mu\nu}, \quad (30)
$$

where $\Gamma^{\lambda}_{\mu\nu}$ is the usual Christoffel three-symbol given in (5), and

$$
\mathcal{W}^{\lambda}_{\mu\nu} = g^{\lambda}{}_{\mu}\mathcal{I}_{\nu} + g^{\lambda}{}_{\nu}\mathcal{I}_{\mu} - g_{\mu\nu}\mathcal{I}^{\lambda} \quad (31)
$$

is the (redefined) Weyl tensor, and the new non-tensorial object:

$$
Q^{\lambda}_{\mu\nu} = g^{\lambda}{}_{\mu}Q_{\nu} + g^{\lambda}{}_{\nu}Q_{\mu} - g_{\mu\nu}Q^{\lambda} \quad (32)
$$

is a new affine connection that is defined thereon the hybrid Riemann-Weyl space and its purpose is to allow the hybrid Riemann-Weyl affine $\overline{\Gamma}^{\lambda}_{\mu\nu}$ to be a tensor. Let us call this affine the Q-affine connection or simply the Q-affine. The geometry that we have just described is what we shall call the *New Weyl Geometry* (NWG).

4.2.5 Transformation of the Riemann-Weyl affine

Now from (6), we know that the old Christoffel three-symbol $\Gamma^{\lambda}_{\mu\nu}$ transforms as follows:

$$
\Gamma^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^{\delta}}\frac{\partial x^{\mu}}{\partial x^{\mu'}}\frac{\partial x^{\nu}}{\partial x^{\nu'}}\Gamma^{\delta}_{\mu\nu} + \frac{\partial x^{\lambda'}}{\partial x^{\delta}}\frac{\partial^{2} x^{\delta}}{\partial x^{\mu'}\partial x^{\nu'}} \quad (33)
$$

and that $\mathcal{W}^{\lambda}_{\mu\nu}$ is a tensor, hence, it transforms like a tensor. Verily, if the Q-affine $Q^{\lambda'}_{\mu'\nu'}$ were to transform just as the Christoffel three-symbol $\Gamma^{\lambda'}_{\mu'\nu'}$, as follows:

$$
Q^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^{\delta}}\frac{\partial x^{\mu}}{\partial x^{\mu'}}\frac{\partial x^{\nu}}{\partial x^{\nu'}}Q^{\delta}_{\mu\nu} + \frac{\partial x^{\lambda'}}{\partial x^{\delta}}\frac{\partial^{2} x^{\delta}}{\partial x^{\mu'}\partial x^{\nu'}}, \quad (34)
$$

then it follows and goes without saying that the object $\overline{\Gamma}^{\lambda}_{\mu\nu}$ will clearly be a tensor. Because Q_{μ} is an affine vector, the Q-tensor will transform as desired, that is, as given in (34), hence the object $\overline{\Gamma}^{\lambda}_{\mu\nu}$ will be a tensor as desired. What this all means is that Q is a pseudo-scalar and not a pure scalar. This is exactly what we did in [33]. Therein [33], we achieved this by forcing Q_{μ} to yield in the favour of our desires and transform as an affine vector as defined in §4.2.2. The resulting theory that one can build from this NWG has been presented in [33] and, in the present paper, it is this same theory that we are now improving on.

As one can verify for themselves, this theory of [33] produces field equations that we are already familiar with – i.e. the Maxwell equations [14]. At this stage of the development of the theory – whether or not the resulting theory is correct – what is important for the reader to appreciate – as has just been here demonstrated thus far – is that a tensorial affine theory can be attained. The problem suffered by Weyl's theory [1] does not apply to the NWG.

5 Theory

We here lay down our theory. What makes the present work different from the preceding works in [8–10] is that the present work incorporates the new results from various research that we have conducted. Because we shall at five instances (i.e. (37), (44), (79a), (79b), and (79c) need to do some gauge fixing, we shall start off by addressing this issue of gauge fixing, i.e. within the context of the present work.

5.1 Gauge fixing

In the physics of *Gauge Theories*, gauge fixing (also called choosing a gauge) denotes a mathematical procedure for coping with redundant degrees of freedom in the field variables. The introduction of a gauge effectively reduces the number of degrees of freedom of the theory. In the present expedition, we shall need the fixing of the gauge and this fixing shall be done in such a manner that one seeks to obtain equations that are congruent with reality. That is, equations that we are already used to know. We shall identify two types of gauges, i.e.:

1. **Natural Gauge:** A *natural gauge* shall here be understood as an exogenous constraint the theory must satisfy in order to meet a global physical requirement. For example, in the present pursuit, we seek a theory based on a spacetime which is such that the magnitude and direction of a vector (tensor) upon parallel transport remains unaltered by the act or procedure of parallel transport of the vector on this spacetime. So, the gauge fixing that will lead us to the attainment of this global symmetry, we shall call a natural gauge or – alternatively – an *exogauge constraint*.

2. **Gauge Constraint:** A *gauge constraint* shall here be understood as an endogenous constraint the theory must satisfy in order to yield equations that are congruent with reality as we are used to know. For example, in the present pursuit, we

seek a theory that will at least yield field equations that are similar to Maxwell's equations [14]. So, the gauge fixing that will lead us to the attainment of such equations, we shall call a gauge constraint or – alternatively – an *endogauge constraint*.

Each time we encounter a natural gauge (exogauge constraint) or a gauge constraint (endogauge constraint), we shall make a clear indication of this.

5.2 Hybrid Riemann-Weyl tensor

From Fig. 1, if say we have a (four-) vector v^λ and we parallel transport it along a closed circuit ABCD in the order (A $\longmapsto$ B) then (B $\longmapsto$ C) then (C $\longmapsto$ D) and then finally (D $\longmapsto$ A), if the space in question has a non-zero curvature, upon arrival at its original location, while the length of this vector may be equal to the length of the original vector, its direction will at the very least be different. The infinitesimal changes of this vector's direction and length along these paths (see e.g. [10, for details of the derivation]), are:

$$dv^\lambda = \overline{R}^\lambda_{\mu\sigma\nu} v^\mu da^\nu db^\sigma \,, \tag{35}$$

where:

$$\overline{R}^\lambda_{\mu\sigma\nu} = \overbrace{\overline{\Gamma}^\lambda_{\mu\nu,\sigma} - \overline{\Gamma}^\lambda_{\mu\sigma,\nu}}^{\text{linear terms}} + \underbrace{\overline{\Gamma}^\lambda_{\delta\sigma}\overline{\Gamma}^\delta_{\mu\nu} - \overline{\Gamma}^\lambda_{\delta\nu}\overline{\Gamma}^\delta_{\mu\sigma}}_{\text{non-linear terms}} \,. \tag{36}$$

is the *Hybrid Riemann-Weyl Tensor*.

5.3 Linear Riemann tensor

Given that we have attained a geometry with tensorial affinities, it goes without saying that – insofar as the beleaguering problems besieging pure Riemann geometry is concerned – now is our time to reap the sweet fruits of our hard labour i.e. it is time to take the fullest advantage of the tensorial nature of the affinities. We now have at our disposal the mathematical and physical prerogative, legitimacy and liberty to choose a spacetime where the non-linear terms do not vanish identically i.e. $\overline{\Gamma}^\lambda_{\mu\nu} \neq 0$, but are bound by the gauge constraint*:

$$\overline{\Gamma}^\delta_{\mu\nu}\overline{\Gamma}^\lambda_{\delta\sigma} = \overline{\Gamma}^\lambda_{\delta\nu}\overline{\Gamma}^\delta_{\mu\sigma} \,. \qquad \textbf{(gauge constraint)} \tag{37}$$

Clearly, from this, the resulting Riemann tensor becomes linear, i.e.:

$$\overline{R}^\lambda_{\mu\sigma\nu} = \overline{\Gamma}^\lambda_{\mu\nu,\sigma} - \overline{\Gamma}^\lambda_{\mu\sigma,\nu} \,. \tag{38}$$

Just like that, we have thrown the non-linear terms out of our sight once and for all-time.

Clearly and without any doubt, this fact that we have chosen a spacetime that is governed by the gauge constraint (37),

means that we have just rid ourselves of the troublesome non-linear terms in the Riemann tensor (38), because with this beautiful and elegant choice (37), the non-linear terms now vanish identically to become but footnotes of history. The justification for this choice of gauge will become clear later when we derive from this tensor (38), the Maxwell equations [14] that we are used to know – albeit this time, these equations are being derived not for the electrodynamic phenomenon, but for the gravitodynamic phenomenon. In the next subsection, we will redefine the Riemann metric $g_{\mu\nu}$ in terms of the four-vector field $\mathcal{A}_\mu$ *via* the decomposition of the metric.

5.4 Decomposition of the metric tensor

A key feature of the present theory, as well as the previous versions of it as given in [8–10], is that of the decomposition of the metric tensor. The Riemann metric $g_{\mu\nu}$ is a compound rank two tensor field symmetric in the $\mu\nu$-indices and because of this symmetry, it consists of ten independent functions. In the present, the components of the metric tensor $\overline{g}_{\mu\nu}$ are a product of the components of a four-vector field $\mathcal{A}_\mu$, thus – this metric consists of four independent functions instead of ten as is the case in pure Riemann geometry.

The covariant $\mathcal{A}_\mu$ and contravariant $\mathcal{A}^\mu$ four-vectors are here to be defined as follows:

$$\mathcal{A}_\mu = (\mathcal{A}^\mu)^\dagger \tag{39}$$

where the dagger-operation ($\dagger$) is the usual *transpose-complex-conjugate operation* applied to the object in question[†], while the covariant $g_{\mu\nu}$, contravariant $g^{\mu\nu}$ and mixed covariant and contravariant metric $g_\mu{}^\nu, g^\mu{}_\nu$ tensors are defined in terms of the covariant $\mathcal{A}_\mu$ and contravariant $\mathcal{A}^\mu$ four-vectors as follows:

$$
\begin{aligned}
g_{\mu\nu} &= \mathcal{A}_\mu \mathcal{A}_\nu \,, & g_\mu{}^\nu &= \mathcal{A}_\mu \mathcal{A}^\nu \,, \\
g^{\mu\nu} &= \mathcal{A}^\mu \mathcal{A}^\nu \,, & g^\mu{}_\nu &= \mathcal{A}^\mu \mathcal{A}_\nu \,.
\end{aligned}
\tag{40}
$$

The mixed covariant and contravariant metric $g_\mu{}^\nu$ and $g^\mu{}_\nu$ tensors are in Riemann defined such in terms of the covariant $g_{\mu\nu}$ and contravariant $g^{\mu\nu}$ as follows:

$$
\begin{aligned}
g_\mu{}^\nu &= g_{\mu\delta}g^{\delta\nu} = \mathcal{A}_\delta \mathcal{A}^\delta g_\mu{}^\nu = g_\delta{}^\delta g_\mu{}^\nu = \delta_\mu{}^\nu \\
g^\mu{}_\nu &= g^{\mu\delta}g_{\delta\nu} = \mathcal{A}^\delta \mathcal{A}_\delta g^\mu{}_\nu = g^\delta{}_\delta g^\mu{}_\nu = \delta^\mu{}_\nu
\end{aligned}
\tag{41}
$$

where $\delta_\mu{}^\nu$ and $\delta^\mu{}_\nu$ are the usual Kronecker-Delta functions. From (41), it follows that:

$$g_\delta{}^\delta = \mathcal{A}_\delta \mathcal{A}^\delta = g^\delta{}_\delta = \mathcal{A}^\delta \mathcal{A}_\delta = 4 \,. \tag{42}$$

---

*This gauge constraint allows us to obtain linear equations. This constraint is made possible by the fact that the affine connections are tensors.

[†]The four-vector $\mathcal{A}_\mu$ can either be a 4×4 or zero rank object. We are not sure at the moment which is which. If it turns out that $\mathcal{A}_\mu$ is a zero rank object, then the *dagger-operation* simple reduces to a *complex-conjugate operation*.

On this new Riemann-Weyl spacetime, the usual raising and lowering of the indices applicable in Riemann geometry holds, i.e.:

$$V_\mu = g_{\mu\delta} V^\delta = g_\mu{}^\delta V_\delta$$
$$V^\mu = g^{\mu\delta} V_\delta = g^\mu{}_\delta V^\delta \qquad (43)$$

With the metric now having been redefined and its nature regarding the lowering and raising of indices, and that the length of the four-vector $\mathcal{A}_\mu$ is four units throughout all spacetime, we will proceed in the next subsection to deduce the first set of the field equations.

5.5 Field equations

Having set the stage, we shall now proceed to write down the resulting field equations.

5.5.1 Field equations I

If both the length and angles are to remain unaltered upon parallel transport, this can only happen if the curvature tensor $\bar{R}^\lambda_{\mu\sigma\nu}$ vanishes at all points of this spacetime, i.e.:

$$\bar{R}^\lambda_{\mu\sigma\nu} = 0. \qquad \textbf{(natural gauge)} \qquad (44)$$

Eq. (44) is a natural equation of the geometry; it emanates from the hypothesis of requiring that both the length and angles are to remain unaltered upon parallel transport. In general, the affine $\bar{\Gamma}^\lambda_{\mu\nu}$ is non-vanishing, i.e. $\bar{\Gamma}^\lambda_{\mu\nu} \neq 0$. So, the present *Hybrid Riemann-Weyl Spacetime* (HRWS) is a curvature-less space because vectors maintain or preserve both their length and orientation under parallel transport. Embedded or cojoined in this HRWS curvature tensor $\bar{R}^\lambda_{\mu\sigma\nu}$ are the Riemann curvature tensor $R^\lambda_{\mu\sigma\nu}$ and the geometrically derived material tensor $\mathcal{T}^\lambda_{\mu\sigma\nu}$. Because of the vanishing nature of HRWS curvature tensor $\bar{R}^\lambda_{\mu\sigma\nu}$, together with its linear nature (see §5.3), we will in the next subsection use these facts to unbundle the Riemann curvature tensor and the material tensor, thereby achieve what Einstein desired but failed to achieve – i.e. a material field derived from pure geometry.

5.5.2 Field equations II

Now that we have a theory linear in the curvature tensor – i.e. a theory in which the non-linear terms vanish – we can use this to separate the Weyl terms $\mathcal{W}^\lambda_{\mu\sigma\nu}$ from the Riemann terms $R^\lambda_{\mu\sigma\nu}$ and as well from the Q-tensor $Q^\lambda_{\mu\sigma\nu}$. That is, we can now rewrite the linear Riemann-Weyl curvature tensor $\bar{R}^\lambda_{\mu\sigma\nu}$ as is given in (38) as follows:

$$\bar{R}^\lambda_{\mu\sigma\nu} = R^\lambda_{\mu\sigma\nu} - \underbrace{\left(\mathcal{W}^\lambda_{\mu\sigma\nu} + Q^\lambda_{\mu\sigma\nu} \right)}_{\mathcal{T}^\lambda_{\mu\sigma\nu}} \qquad (45)$$

where:

$$R^\lambda_{\mu\sigma\nu} = \Gamma^\lambda_{\mu\nu,\sigma} - \Gamma^\lambda_{\mu\sigma,\nu} \qquad \text{(a)}$$
$$\mathcal{W}^\lambda_{\mu\sigma\nu} = \mathcal{W}^\lambda_{\mu\nu,\sigma} - \mathcal{W}^\lambda_{\mu\sigma,\nu} \qquad \text{(b)} \qquad (46)$$
$$Q^\lambda_{\mu\sigma\nu} = Q^\lambda_{\mu\nu,\sigma} - Q^\lambda_{\mu\sigma,\nu} \qquad \text{(c)}$$

are the *linear Riemann curvature tensor* (46a), the *linear Weyl curvature tensor* (46b), and the *linear Q-curvature tensor* (46c) or simply the *Q-tensor*.

An excogitative inspection of the Riemann curvature tensor will clearly reveal that this tensor is a function of the four-vector field $\mathcal{A}_\mu$, i.e. $R^\lambda_{\mu\sigma\nu} = R^\lambda_{\mu\sigma\nu}(\mathcal{A}_\alpha)$, while the Weyl and the Q-tensors are functions of ϱ, i.e. $\mathcal{W}^\lambda_{\mu\sigma\nu} = \mathcal{W}^\lambda_{\mu\sigma\nu}(\mathcal{I}_\alpha)$ and $Q^\lambda_{\mu\sigma\nu} = Q^\lambda_{\mu\sigma\nu}(\varrho)$. The Q-tensor is a direct function of ϱ while the Weyl tensor is not – remember (25) that $\mathcal{I}_\alpha = -\frac{1}{2}\partial_\alpha \varrho$, hence, as said $\mathcal{W}^\lambda_{\mu\sigma\nu} = \mathcal{W}^\lambda_{\mu\sigma\nu}(\varrho)$. Why are we talking of the functional dependence of these tensors?

The reason for excogitating on the functional dependence of these tensors is that we not only want to, but shall identify the Riemann curvature tensor as describing *Einstein's beautiful marble* that, in Einstein's vision and desideratum, is described by the metric tensor $g_{\mu\nu}$; while the Weyl curvature tensor and the Q-curvature tensor describe *Einstein's ugly wood* – albeit – varnished (polished) wood this time around since the field ϱ is later to be identified with the beautiful – albeit – arcane quantum mechanical object, namely the quantum probability amplitude.

After the above deliberations, it therefore makes much sense to house the Weyl curvature tensor and the Q-curvature tensor under one roof since they constitute the material tensor. To that end, let us represent the sum total material curvature tensor using the symbol $\mathcal{T}^\lambda_{\mu\sigma\nu}$ where:

$$\mathcal{T}^\lambda_{\mu\sigma\nu} = \mathcal{W}^\lambda_{\mu\sigma\nu} + Q^\lambda_{\mu\sigma\nu}. \qquad (47)$$

With the above definition (47) of the material tensor, it follows that the Riemann-Weyl curvature tensor $\bar{R}^\lambda_{\mu\sigma\nu}$ can now be written as an object comprising two main tensors expressing the fields ($R^\lambda_{\mu\sigma\nu}$) and their corresponding material ($\mathcal{T}^\lambda_{\mu\sigma\nu}$) counterpart:

$$\bar{R}^\lambda_{\mu\sigma\nu} = R^\lambda_{\mu\sigma\nu} - \mathcal{T}^\lambda_{\mu\sigma\nu}. \qquad (48)$$

What we have done – from (45) to (48) above – is to indulge and cajole the reader to the idea of envisioning the Riemann-Weyl tensor in Einstein's vision of a *marble* and *wood* component, albeit, with the wood now recast into its quantum mechanical description.

Now, from (44) and (48), it follows that:

$$R^\lambda_{\mu\sigma\nu} = \mathcal{T}^\lambda_{\mu\sigma\nu}. \qquad (49)$$

At this point – if it turns out that this theory proves to be a correct description of physical and natural reality as we know it – we have no doubt in our mind that if Einstein were watching from above or from wherever in the interstices of spacetime,

he must be smiling endlessly because his lifelong endeavour was to derive[*] the material tensor from pure geometry and not to insert it by sleight of mind as he did with his gravitational field (7). In-line with Einstein's deepest quest and longing insofar in attaining a final UFT of all the forces of Nature, we have in the present derived the material tensor from pure geometry.

As we saw previously in §2.5, Einstein's ultimate goal was to turn *wood* into *marble* so to speak, which meant deriving the material field from pure geometry. Einstein wanted to find the final theory; this he pursued to the very end of his life to a point that while on his deathbed on April 18, 1955, instead of worrying about the imminent end of his fruitful life, he asked for a pen and his notes so that he could continue to work on the unified field theory that he was working on at the time. It is sad to say that Einstein never laid a fertile egg on this front – i.e. the front of unification.

Be that as it may, it is without an iota of doubt that we say that if what is before us proves itself to have a correspondence with physical and natural reality, then we can safely say we have achieved one of Einstein's goals to attaining the "elicit dream of a Final Theory" by deriving the material tensor from pure geometry – wood, one way or the other, has finally been turned into marble! This we are certain has been achieved in the present UFT. The only question is, *Does the theory correspond with physical and natural reality?* This we leave in the able hands of our reader so that they may be their own judge on that very important matter.

5.5.3 Field equations III

First Voss-Bianchi Identities: Further, we shall derive other field equations. We know that the Riemann curvature tensor satisfies the first Voss-Bianchi[†] identity, namely:

$$R^\lambda_{\mu\sigma\nu} + R^\lambda_{\nu\mu\sigma} + R^\lambda_{\sigma\nu\mu} \equiv 0 \,. \tag{50}$$

From this first Bianchi identity and as well from (49), it follows that:

$$\mathcal{T}^\lambda_{\mu\sigma\nu} + \mathcal{T}^\lambda_{\nu\mu\sigma} + \mathcal{T}^\lambda_{\sigma\nu\mu} \equiv 0 \,. \tag{51}$$

In the next subsection, we present the second Voss-Bianchi identity.

---

[*]Here, we must hasten to say that we have not exactly derived the material tensor field $\mathcal{T}^\lambda_{\mu\sigma\nu}$, but merely justified its physical existence on the fundamental basis of the need for tensorial affinities. Thus, this material field is not only justifiable on a fundamental physical level, but very much a part and parcel of the whole edifice of the marvellous structure of the spacetime continuum.

[†]In the wider literature – if not every common text where these identities are considered – they are referred to as the *Bianchi Identities* after the Italian mathematician – Luigi Bianchi (1856-1928) who published them in 1902 [60]. However, the reality to the matter is that these identities were first derived and published by the German mathematician Aurel Voss (1845-1931) in 1880 [61]. Hence, keeping matters in their correct historic record and perspective, and to give due credit and acknowledgement of the work of Aurel Voss, we herein refer to these identities ((50) and (52)) as the *Voss-Bianchi Identities*.

5.5.4 Field equations IV

Second Voss-Bianchi Identities: Furthermore, we are going to derive our last set of field equations. We know that the Riemann curvature tensor satisfies the second Voss-Bianchi identity, namely:

$$R^\lambda_{\upsilon\mu\sigma,\nu} + R^\lambda_{\upsilon\nu\mu,\sigma} + R^\lambda_{\upsilon\sigma\nu,\mu} \equiv 0 \,. \tag{52}$$

From this second Bianchi identity and as well from (49), it follows that:

$$\mathcal{T}^\lambda_{\upsilon\mu\sigma,\nu} + \mathcal{T}^\lambda_{\upsilon\nu\mu,\sigma} + \mathcal{T}^\lambda_{\upsilon\sigma\nu,\mu} \equiv 0 \,. \tag{53}$$

In the next section, we shall explore (49), (50), (51), (52) and (53), and from these equations, we shall see that one is able to obtain field equations that we are already familiar with. Before we depart this section, we must say that while we have shown that the material tensor $\mathcal{T}^\lambda_{\upsilon\mu\sigma,\nu}$ does satisfy the Voss-Bianchi identities, the subcomponents $(\mathcal{W}^\lambda_{\upsilon\mu\sigma,\nu}; Q^\lambda_{\upsilon\mu\sigma,\nu})$ of this tensor also satisfy the Voss-Bianchi identities, i.e.:

$$
\begin{aligned}
\mathcal{W}^\lambda_{\mu\sigma\nu} + \mathcal{W}^\lambda_{\nu\mu\sigma} + \mathcal{W}^\lambda_{\sigma\nu\mu} &\equiv 0 & \text{(a)} \\
Q^\lambda_{\mu\sigma\nu} + Q^\lambda_{\nu\mu\sigma} + Q^\lambda_{\sigma\nu\mu} &\equiv 0 & \text{(b)} \\
\mathcal{W}^\lambda_{\upsilon\mu\sigma,\nu} + \mathcal{W}^\lambda_{\upsilon\nu\mu,\sigma} + \mathcal{W}^\lambda_{\upsilon\sigma\nu,\mu} &\equiv 0 & \text{(c)} \\
Q^\lambda_{\upsilon\mu\sigma,\nu} + Q^\lambda_{\upsilon\nu\mu,\sigma} + Q^\lambda_{\upsilon\sigma\nu,\mu} &\equiv 0 & \text{(d)}
\end{aligned}
\tag{54}
$$

where in (54a,b) and (54c,d), we have the first and second Voss-Bianchi identities of $\mathcal{W}^\lambda_{\upsilon\mu\sigma,\nu}$ and $Q^\lambda_{\upsilon\mu\sigma,\nu}$ respectively.

6 Affine, Riemann and the material tensor

In the present section, we are going to calculate or express the affine tensor $\Gamma^\lambda_{\mu\nu}$, the Riemann tensor $R_{\mu\nu}$, and the material tensor $\mathcal{T}_{\mu\nu}$ in terms of a Maxwell field tensor $\mathfrak{F}_{\mu\nu}$. This exercise is meant to prepare us for the work to be conducted in §7 where we are going to write down our desired *Maxwell Gravitomagnetic Field Equations*.

6.1 Affine tensor

We already know from (5) that the affine connection $\Gamma^\lambda_{\mu\nu}$ is such that $2\Gamma^\lambda_{\mu\nu} = g^{\delta\lambda}\left(g_{\delta\mu,\nu} + g_{\nu\delta,\mu} - g_{\mu\nu,\delta}\right)$, and from the present new findings that the decomposed Riemann metric tensor is such that $g_{\mu\nu} = \mathcal{A}_\mu \mathcal{A}_\nu$. What we want – and will – do here is to substitute the decomposed metric into the affine wherefrom we expect to obtain the usual Maxwell-type field tensor of electromagnetism. To that end, we substitute the metric into the affine and then differentiate this metric as required by the differentials in the affine – doing so, we obtain:

$$2\Gamma^\lambda_{\mu\nu} = g^{\delta\lambda}\Big[\overbrace{\mathcal{A}_\delta \mathcal{A}_{\mu,\nu}}^{\text{Term I}} + \overbrace{\mathcal{A}_\mu \mathcal{A}_{\delta,\nu}}^{\text{Term II}} + \overbrace{\mathcal{A}_\delta \mathcal{A}_{\nu,\mu}}^{\text{Term III}} + \underbrace{\mathcal{A}_\nu \mathcal{A}_{\delta,\mu}}_{\text{Term IV}} - \underbrace{\mathcal{A}_\mu \mathcal{A}_{\nu,\delta}}_{\text{Term V}} - \underbrace{\mathcal{A}_\nu \mathcal{A}_{\mu,\delta}}_{\text{Term VI}} \Big] \,. \tag{55}$$

Now, we shall identify the labelled terms in (55), that is, terms that will yield for us the desired Maxwell-type field tensor of electromagnetism.

1. **Terms II and V:** Combining Term II and Term V, we will have:

$$\mathcal{A}_\mu \mathfrak{F}_{\delta\nu} = \mathcal{A}_\mu \left(\mathcal{A}_{\delta,\nu} - \mathcal{A}_{\nu,\delta} \right) \tag{56}$$

where:

$$\mathfrak{F}_{\delta\nu} = \mathcal{A}_{\delta,\nu} - \mathcal{A}_{\nu,\delta} \tag{57}$$

is the gravitomagnetic field tensor. This tensor (57) is our desired Maxwell-type field tensor of electromagnetism – albeit – this time – as per our desire – it is appearing in the equations of gravitation and not electromagnetism.

2. **Terms IV and VI:** Further, combining Term IV and Term VI, we will have:

$$\mathcal{A}_\nu \mathfrak{F}_{\delta\mu} = \mathcal{A}_\nu \left(\mathcal{A}_{\delta,\mu} - \mathcal{A}_{\mu,\delta} \right) \tag{58}$$

where – as in (57):

$$\mathfrak{F}_{\delta\mu} = \mathcal{A}_{\delta,\mu} - \mathcal{A}_{\mu,\delta} \tag{59}$$

is the same gravitomagnetic field tensor – the only difference is the interchange of the indices.

3. **Terms I and III:** Lastly, combining Term I and Term III, we will have:

$$\mathcal{A}_\delta \Omega_{\mu\nu} = \mathcal{A}_\delta \left(\mathcal{A}_{\mu,\nu} + \mathcal{A}_{\nu,\mu} \right) \tag{60}$$

where – this time:

$$\Omega_{\mu\nu} = \mathcal{A}_{\mu,\nu} + \mathcal{A}_{\nu,\mu} \tag{61}$$

is not a gravitomagnetic field tensor, but some non-tensorial object that will prove to be absolutely essential and necessary in the generation of the source-free Maxwell-type equations for gravitomagnetism.

From the foregoing, it follows from (57), (59) and (61), that:

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\delta\lambda} \left[\mathcal{A}_\mu \mathfrak{F}_{\delta\nu} + \mathcal{A}_\nu \mathfrak{F}_{\delta\mu} + \mathcal{A}_\delta \Omega_{\mu\nu} \right]. \tag{62}$$

Now, multiplying the terms in the square bracket by $g^{\delta\lambda}$, the meaning of which is that we have to raise the δ-index in these square brackets and reset it so that it now equals λ, i.e.:

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} \left[\mathcal{A}_\mu \mathfrak{F}^\lambda_{\ \nu} + \mathcal{A}_\nu \mathfrak{F}^\lambda_{\ \mu} + \mathcal{A}^\lambda \Omega_{\mu\nu} \right]. \tag{63}$$

In (63), we most importantly have expressed the Christoffel affine in terms of the Maxwell field tensor $\mathfrak{F}_{\mu\nu}$. In the next section, we shall proceed to express the Riemann tensor in terms of the same Maxwell field tensor $\mathfrak{F}_{\mu\nu}$.

For the purposes of convenience in the coming computations to be made in the subsequent sections, we shall write down the Christoffel affine (i.e. (63)), as follows:

$$\Gamma^\lambda_{\mu\nu} = \check{\Gamma}^\lambda_{\mu\nu} + \Omega^\lambda_{\mu\nu} \tag{64}$$

where:

$$\check{\Gamma}^\lambda_{\mu\nu} = \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda_{\ \nu} + \mathcal{A}_\nu \mathfrak{F}^\lambda_{\ \mu} \right) \tag{65}$$

and:

$$\Omega^\lambda_{\mu\nu} = \frac{1}{2} \mathcal{A}^\lambda \Omega_{\mu\nu}. \tag{66}$$

The object $\check{\Gamma}^\lambda_{\mu\nu}$ is a tensor while $\Omega^\lambda_{\mu\nu}$ is not, for, upon a transformation of the system of coordinates, this affine $\Omega^\lambda_{\mu\nu}$ transforms in the exact same manner as the Christoffel symbols (see (6)), that is, it transforms as follows:

$$\Omega^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^\delta} \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \Omega^\delta_{\mu\nu} + \frac{\partial x^{\lambda'}}{\partial x^\delta} \frac{\partial^2 x^\delta}{\partial x^{\mu'} \partial x^{\nu'}}. \tag{67}$$

In the next subsection, as we continue to work toward the writing down of the resultant field equations, we shall express the Riemann tensor in terms of the gravitomagnetic Maxwell-type tensor $\mathfrak{F}_{\mu\nu}$.

6.2 Riemann tensor

We are not only going to express the Riemann tensor in terms of the gravitomagnetic Maxwell-type field tensor $\mathfrak{F}_{\mu\nu}$ but decompose this tensor into three tensors. To that end, we will start-off by substituting the newly re-expressed Christoffel affine in (64) into the linear Riemann tensor (46a); so doing, we obtain:

$$R^\lambda_{\mu\sigma\nu} = \check{\Gamma}^\lambda_{\mu\nu,\sigma} - \check{\Gamma}^\lambda_{\mu\sigma,\nu} + \Omega^\lambda_{\mu\nu,\sigma} - \Omega^\lambda_{\mu\sigma,\nu}$$
$$= \check{R}^\lambda_{\mu\sigma\nu} + \Omega^\lambda_{\mu\sigma\nu} \tag{68}$$

where:

$$\check{R}^\lambda_{\mu\sigma\nu} = \check{\Gamma}^\lambda_{\mu\nu,\sigma} - \check{\Gamma}^\lambda_{\mu\sigma,\nu}$$
$$\Omega^\lambda_{\mu\sigma\nu} = \Omega^\lambda_{\mu\nu,\sigma} - \Omega^\lambda_{\mu\sigma,\nu} \tag{69}$$

are tensors. The reader will need to verify for themselves that – indeed – these objects are tensors.

Further, we will express $\check{R}^\lambda_{\mu\sigma\nu}$ in terms of the field tensor $\mathfrak{F}_{\mu\nu}$ by substituting $\check{\Gamma}^\lambda_{\mu\nu}$ as it is given in (65); so doing, one obtains:

$$\check{R}^\lambda_{\mu\sigma\nu} = \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda_{\ \nu,\sigma} + \mathcal{A}_\nu \mathfrak{F}^\lambda_{\ \mu,\sigma} \right) -$$
$$- \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda_{\ \sigma,\nu} + \mathcal{A}_\sigma \mathfrak{F}^\lambda_{\ \mu,\nu} \right) +$$
$$+ \frac{1}{2} \left(\mathcal{A}_{\mu,\sigma} \mathfrak{F}^\lambda_{\ \nu} + \mathcal{A}_{\nu,\sigma} \mathfrak{F}^\lambda_{\ \mu} \right) -$$
$$- \frac{1}{2} \left(\mathcal{A}_{\mu,\nu} \mathfrak{F}^\lambda_{\ \sigma} + \mathcal{A}_{\sigma,\nu} \mathfrak{F}^\lambda_{\ \mu} \right)$$
$$= \acute{R}^\lambda_{\mu\sigma\nu} + \grave{R}^\lambda_{\mu\sigma\nu} \tag{70}$$

where:

$$\acute{R}^\lambda_{\mu\sigma\nu} = \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda_{\ \nu,\sigma} + \mathcal{A}_\nu \mathfrak{F}^\lambda_{\ \mu,\sigma} \right)$$
$$- \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda_{\ \sigma,\nu} + \mathcal{A}_\sigma \mathfrak{F}^\lambda_{\ \mu,\nu} \right) \tag{71}$$

and:

$$\grave{R}^\lambda_{\mu\sigma\nu} = \frac{1}{2} \left(\mathcal{A}_{\mu,\sigma} \mathfrak{F}^\lambda_{\ \nu} + \mathcal{A}_{\nu,\sigma} \mathfrak{F}^\lambda_{\ \mu} \right) -$$
$$- \frac{1}{2} \left(\mathcal{A}_{\mu,\nu} \mathfrak{F}^\lambda_{\ \sigma} + \mathcal{A}_{\sigma,\nu} \mathfrak{F}^\lambda_{\ \mu} \right) \tag{72}$$

are tensors. Once again, the reader will need to verify for themselves that these objects are indeed tensors. Therefore, from (68) and (70), it follows that:

$$R^\lambda_{\mu\sigma\nu} = \acute{R}^\lambda_{\mu\sigma\nu} + \grave{R}^\lambda_{\mu\sigma\nu} + \Omega^\lambda_{\mu\sigma\nu}. \tag{73}$$

In (73), we have – as desired – not only re-expressed the Riemann tensor, but decomposed it into three part tensors. Now – in the next subsection, we will conduct the same exercise with the material tensor. All this re-expression and decomposition is all gearing up for the derivation of the result field equation of the theory.

6.3 Material tensor

Just as we have decomposed the Riemann curvature tensor into three parts in (73), we are now going to decompose the material curvature tensor $\mathcal{T}^\lambda_{\mu\sigma\nu}$ into three parts by decomposing into two parts, the linear Weyl curvature tensor $\mathcal{W}^\lambda_{\mu\sigma\nu}$. To that end, decomposing the Weyl part of the material tensor field by differentiating the products $g^\lambda_\mu \mathcal{I}_\nu$, we obtain that:

$$\begin{aligned}
\mathcal{T}^\lambda_{\mu\sigma\nu} = & \left(g^\lambda_\mu \mathcal{I}_{\nu,\sigma} + g_\nu{}^\lambda \mathcal{I}_{\mu,\sigma} - g_{\mu\nu}\mathcal{I}^\lambda{}_{,\sigma}\right) - \\
& - \left(g^\lambda_\mu \mathcal{I}_{\sigma,\nu} + g_\sigma{}^\lambda \mathcal{I}_{\mu,\nu} - g_{\mu\sigma}\mathcal{I}^\lambda{}_{,\nu}\right) + \\
& + \left(g^\lambda_{\mu,\sigma} \mathcal{I}_\nu + g_{\nu,\sigma}{}^\lambda \mathcal{I}_\mu - g_{\mu\nu,\sigma}\mathcal{I}^\lambda\right) - \\
& - \left(g^\lambda_{\mu,\nu} \mathcal{I}_\sigma + g_{\sigma,\nu}{}^\lambda \mathcal{I}_\mu - g_{\mu\sigma,\nu}\mathcal{I}^\lambda\right) + \\
& + Q^\lambda_{\mu\sigma\nu} \\
= & \acute{\mathcal{T}}^\lambda_{\mu\sigma\nu} + \grave{\mathcal{T}}^\lambda_{\mu\sigma\nu} + Q^\lambda_{\mu\sigma\nu}
\end{aligned} \tag{74}$$

where the newly introduced tensors $\acute{\mathcal{T}}^\lambda_{\mu\sigma\nu}$ and $\grave{\mathcal{T}}^\lambda_{\mu\sigma\nu}$ are explicitly defined as follows:

$$\begin{aligned}
\acute{\mathcal{T}}^\lambda_{\mu\sigma\nu} = & \left(g^\lambda_\mu \mathcal{I}_{\nu,\sigma} + g_\nu{}^\lambda \mathcal{I}_{\mu,\sigma} - g_{\mu\nu}\mathcal{I}^\lambda{}_{,\sigma}\right) - \\
& - \left(g^\lambda_\mu \mathcal{I}_{\sigma,\nu} + g_\sigma{}^\lambda \mathcal{I}_{\mu,\nu} - g_{\mu\sigma}\mathcal{I}^\lambda{}_{,\nu}\right) \\
= & \left[g_\nu{}^\lambda \mathcal{I}_{\mu,\sigma} - g_{\mu\nu}\mathcal{I}^\lambda{}_{,\sigma}\right] - \left[g_\sigma{}^\lambda \mathcal{I}_{\mu,\nu} - g_{\mu\sigma}\mathcal{I}^\lambda{}_{,\nu}\right]
\end{aligned} \tag{75}$$

and:

$$\begin{aligned}
\grave{\mathcal{T}}^\lambda_{\mu\sigma\nu} = & \left(g^\lambda_{\mu,\sigma} \mathcal{I}_\nu + g_{\nu,\sigma}{}^\lambda \mathcal{I}_\mu - g_{\mu\nu,\sigma}\mathcal{I}^\lambda\right) - \\
& - \left(g^\lambda_{\mu,\nu} \mathcal{I}_\sigma + g_{\sigma,\nu}{}^\lambda \mathcal{I}_\mu - g_{\mu\sigma,\nu}\mathcal{I}^\lambda\right).
\end{aligned} \tag{76}$$

Written in a much clearer manner:

$$\mathcal{T}^\lambda_{\mu\sigma\nu} = \acute{\mathcal{T}}^\lambda_{\mu\sigma\nu} + \grave{\mathcal{T}}^\lambda_{\mu\sigma\nu} + Q^\lambda_{\mu\sigma\nu}. \tag{77}$$

At this juncture, having now written down the Riemann and the material curvature tensors in the manner that we have written them in (73) and (77), we are now ready to explore the *Resultant Field Equations*.

7 Resultant field equations

Having calculated in (73) and (77), the Riemann and the material curvature tensors into a form that allows us to execute the main business of the day of deriving (deducing) the source-coupled and source-free field equations respectively, we are going to start by writing main field (49) with the decoupled Riemann and the material curvature tensors, i.e.:

$$\overbrace{\underbrace{\acute{R}^\lambda_{\mu\sigma\nu}}_{\text{L I}} + \underbrace{\grave{R}^\lambda_{\mu\sigma\nu}}_{\text{L II}} + \underbrace{\Omega^\lambda_{\mu\sigma\nu}}_{\text{L III}}}^{\text{marble}} = \overbrace{\underbrace{\acute{\mathcal{T}}^\lambda_{\mu\sigma\nu}}_{\text{R I}} + \underbrace{\grave{\mathcal{T}}^\lambda_{\mu\sigma\nu}}_{\text{R II}} + \underbrace{Q^\lambda_{\mu\sigma\nu}}_{\text{R III}}}^{\text{varnished wood}}. \tag{78}$$

$$\underbrace{\hphantom{\acute{R}^\lambda_{\mu\sigma\nu} + \grave{R}^\lambda_{\mu\sigma\nu} + \Omega^\lambda_{\mu\sigma\nu}}}_{R^\lambda_{\mu\sigma\nu}(\mathcal{A}_\alpha)} \quad \underbrace{\hphantom{\acute{\mathcal{T}}^\lambda_{\mu\sigma\nu} + \grave{\mathcal{T}}^\lambda_{\mu\sigma\nu} + Q^\lambda_{\mu\sigma\nu}}}_{\mathcal{T}^\lambda_{\mu\sigma\nu}(\varrho)}$$

Eq. (78) is the single most important equation of our theory and it is out of this equation that we are to derive the rest of the field equations of the theory. The setting up of the said field equations of the theory we shall do by way of introduction of the appropriate gauge constraints. If it were us creating the Universe out of (78), how were we going to proceed to accomplish this monumental task? Our thinking is that a term on the left-hand side in (78) has a corresponding term on the right. Therefore, if our said thinking is reasonable or correct, then our task to finding the sought-for field equations is simply to correctly match the left- and right-hand side terms in (78). If the choice we make turns out to describe our Universe as we know it, then this choice will somehow be the choice that has been made in creating the Universe! This should give us a foothold in seeking answers to some of Einstein's deep philosophical questions about the creation of the Universe.

With regard to the creation of the Universe, Einstein is famously quoted as having said *I want to know the mind of God ... whether or not He had a choice in making the Universe* and on a different occasion, as having said *When I am judging a theory, I ask myself whether, if I were God, I would have arranged the World in such a way.* [62]. These are very deep questions that Einstein was asking about physical and natural reality. Using Einstein's words as a source of inspiration, strength and guidance, we find ourself asking *How are we to construct the resulting field equations from (78)?*

It is with great equanimity that we say that we are of the veritable standpoint that the first term (labelled L I) on the left-hand side of (78) corresponds to the first term on the right-hand side (labelled R I); that, the second term on the left (labelled L II) corresponds to the second term on the right-hand side (labelled R II); and, likewise, that, the L III term corresponds to the R III term, i.e.:

$$\begin{aligned}
\acute{R}^\lambda_{\mu\sigma\nu} &= \acute{\mathcal{T}}^\lambda_{\mu\sigma\nu} & \text{(a)} \\
\grave{R}^\lambda_{\mu\sigma\nu} &= \grave{\mathcal{T}}^\lambda_{\mu\sigma\nu} & \text{(b)} \\
\Omega^\lambda_{\mu\sigma\nu} &= Q^\lambda_{\mu\sigma\nu} & \text{(c)}
\end{aligned} \tag{79}$$

(gauge constraints)

Eqs. (79a), (79b) and (79c) are constraints on (78), albeit endogauge constraints of the theory. Shortly in §7.1 and §7.2, we shall show that (79a) and (79c) are the gravitational source-coupled and source-free Maxwell's field equations [14]. Exploration of (79b) is left for a later paper.

7.1 Source-coupled field equations

As claimed above, we shall now proceed to show that (79a) is indeed the gravitomagnetic Maxwell-type source-coupled field equation. To see this, we shall multiply (79a) on both sides by $\mathcal{A}^{\alpha}$ and thereafter contracting the (α, μ) and (λ, σ)-indices by setting $\alpha = \mu = \beta$ and $\lambda = \sigma = \delta$; so doing, we obtain:

$$\mathcal{A}^{\beta} \acute{R}^{\delta}_{\beta\delta\nu} = \mathcal{A}^{\beta} \acute{\mathcal{T}}^{\delta}_{\beta\delta\nu} . \tag{80}$$

On the other hand, for $\mathcal{A}^{\beta} \acute{R}^{\delta}_{\beta\delta\nu}$, we have that:

$$\mathcal{A}^{\beta} \acute{R}^{\delta}_{\beta\delta\nu} = \mathfrak{F}^{\delta}_{\ \nu,\delta} , \tag{81}$$

and this already looks very familiar – is this not the well known left-hand side of Maxwell's source-coupled field equation [14] – albeit – in the realm of the gravitational phenomenon? It certainly is.

For $\mathcal{A}^{\beta} \acute{\mathcal{T}}^{\delta}_{\beta\delta\nu}$, we have that:

$$\begin{aligned} \mathcal{A}^{\beta} \acute{\mathcal{T}}^{\delta}_{\beta\delta\nu} &= -2\mathcal{A}^{\delta} \mathcal{I}_{\delta,\nu} - \mathcal{I}^{\delta}_{\ ,\delta} \mathcal{A}_{\nu} \\ &= -2\mathcal{A}^{\delta} \partial_{\delta} \partial_{\nu}\varrho + (\Box \varrho / 2)\, \mathcal{A}_{\nu} \\ &= -\widetilde{\mu} \mathfrak{J}_{\nu} + \kappa^2 \mathcal{A}_{\nu} \end{aligned} \tag{82}$$

where from our foreknowledge and, by way of inference and inspiration from experience, we have set in (82):

$$2\mathcal{A}^{\delta} \partial_{\delta} \partial_{\nu}\varrho \ = \ \widetilde{\mu} \mathfrak{J}_{\nu} ,$$

with $\widetilde{\mu}$ being a coupling constant that restores dimensional consistency and $\mathfrak{J}_{\nu}$ is the conserved gravitational four-current density (or four-momentum density). Thus from the foregoing, it follows that $\mathfrak{F}^{\delta}_{\ \nu,\delta} = -\widetilde{\mu} \mathfrak{J}_{\nu}$. We expect that $\widetilde{\mu}$ should embody (represent) Newton's gravitational constant. For aesthetic reasons, we prefer to write this equation $\mathfrak{F}^{\delta}_{\ \nu,\delta} = -\widetilde{\mu} \mathfrak{J}_{\nu}$ in the form:

$$\partial^{\mu} \mathfrak{F}_{\mu\nu} = -\widetilde{\mu} \mathfrak{J}_{\nu} + \kappa^2 \mathcal{A}_{\nu} . \tag{83}$$

The above (83) is Maxwell's source-coupled field equations [14], albeit in the present case, these equations are emerging not in the realm and domain of electrodynamics, but pure gravitation. This derivation of (83) completes the first part of the main task of the present paper. In the next section, we tackle the second part where we shall derive the source-free gravitomagnetic field equations.

7.2 Source-free field equations

Having derived the source-coupled field (83), we are now going to deduce (derive) the source-free field equations from the field (79c) by means of the first Voss-Bianchi identities (in (50)). To that end, we shall achieve this by conducting a cyclic permutation of the $\mu\sigma\nu$-indices in (79c), i.e.:

$$\Omega^{\lambda}_{[\mu\sigma\nu]} = Q^{\lambda}_{[\mu\sigma\nu]} . \tag{84}$$

The square-brackets in (84) here and after indicate the cyclic permutation of the indices for the particular tensor in question.

Now for $Q^{\lambda}_{[\mu\sigma\nu]}$, we already know from (54b) that $Q^{\lambda}_{[\mu\sigma\nu]} \equiv 0$. For $\Omega^{\lambda}_{[\mu\sigma\nu]}$, a computation of this tensor will yield $\Omega^{\lambda}_{[\mu\sigma\nu]} = \mathcal{A}^{\lambda} \mathfrak{F}_{\mu\sigma,\nu} + \mathcal{A}^{\lambda} \mathfrak{F}_{\nu\mu,\sigma} + \mathcal{A}^{\lambda} \mathfrak{F}_{\sigma\nu,\mu}$. Therefore, combining this with (54b) and (84), it follows that:

$$\partial_{\nu} \mathfrak{F}_{\mu\sigma} + \partial_{\sigma} \mathfrak{F}_{\nu\mu} + \partial_{\mu} \mathfrak{F}_{\sigma\nu} \equiv 0 . \tag{85}$$

If anything, the above (85) is indeed Maxwell's source-free field equations [14] written in terms of the covariant derivative, albeit in the present case, this equation is emerging deep within the full domains of gravitation, i.e. from the pure soils of geometry. The derivation of (85) technically completes the main task of the present paper. We surely have shown that one can derive Maxwell's equations [14] from the viewpoint of a Riemann-Weyl geometry standpoint. This must give a strong leverage and impetus to gravitomagnetism as a legitimate and plausible fundamental phenomenon lying well within the domain and realm of real science that is well worthy of the attention of a knowledge seeking scientific mind.

8 Discussion

For what we wanted to achieve in the present paper, we are of the view that we have succeeded – i.e. succeeded in demonstrating that – a legitimate fundamental geometrodynamic justification of gravitomagnetism can be found from the fertile soils of Weyl's [1] beautiful but now thought to be dead and obsolete theory. We further believe that this justification adds much greater impetus to the justification one obtains from say Heras's [43] insightful and powerful existence theorem, or from Behera's [32] interesting theorem that much like the electromagnetic force, the gravitational force is susceptible to a four-vector description. Furthermore, we are also confident that what we have presented herein is being presented for the first time in the scientific literature, hence, these are new blossoms in the realm of ideas.

In the following subsections (i.e. §8.1 and §8.3), we shall discuss (in §8.2) rather briefly, the gauge conditions arising in the present theory and in §8.3, our thoughts regarding a *Quantum Theory of Gravity*. No tangible conclusion is drawn from this paper as this is left for our able and agile reader to makeup their own mind regarding what has herein been presented. We are of the view that this paper is clear and straight forward enough, so much that it should not be difficult to come to a conclusion as to what this paper really means regarding gravitomagnetism.

8.1 Architecture and design of theory

We have used Weyl's modified theory [1] to give a legal and fundamental basis for the existence of gravitomagnetism, and this gravitomagnetic theory can and will be extended in the next paper to demonstrate a possible unity between gravitation and electricity. Naturally and with justification, one will (or may) ask the interesting question: *What in the present have we now done differently that no one has done in the past to this 102 year old theory that suffered a moumental still-birth under the able hands and agile eyes of Albert Einstein's razor sharp intellect whose criticism made sure that Weyl's theory [1] failed?*

In a nutshell, what we have done in our quest to give a fundamental geometrodynamic justification of gravitomagnetism, is to modify Weyl's [1] supposedly failed geometry whose endeavour was to bring the gravitational and electromagnetic forces into one grand scheme, *via* the subtle addition of a conformal scalar leading to the addition of a tensorial affine connection that is a function of a four-vector field and have turned Weyl's [1] scalar into a pseudo-scalar of the second kind. Succinctly stated – in just nine major steps – this is what we have done:

1. The first insight has been to make the Weyl [1] conformal scalar a pseudo-scalar of the second kind and this allows us to obtain tensorial affinities within the realm of Weyl's theory [1].

2. The second insight is to realize that the Riemann metric tensor $g_{\mu\nu}$ can be decomposed into a product of a four-vector $\mathcal{A}_\mu$ so that, instead of describing the metric using ten potentials, it is now described by only four potentials: $g_{\mu\nu} = \mathcal{A}_\mu \mathcal{A}_\nu$.

3. Third – in a Weyl [1] fashion – *via* the newly introduced pseudo-scalar, we added a new non-tensorial affine connection $Q^\lambda_{\mu\nu}$ (i.e. $\bar{\Gamma}^\lambda_{\mu\nu} = \Gamma^\lambda_{\mu\nu} - \mathcal{W}^\lambda_{\mu\nu} - Q^\lambda_{\mu\nu}$) and demanded of it to yield for us a resultant affine connection that is a tensor. Once we have a tensorial affine connections, it means we now have the tool required to obtain Einstein's desired geometry that is such that both the length and direction of a vector under parallel transport are preserved.

4. Fourth, the preservation of both the direction and length of the vector under parallel transport automatically implies that the curvature tensor $\bar{R}^\lambda_{\mu\sigma\nu}$ will vanish identically everywhere, i.e. $\bar{R}^\lambda_{\mu\sigma\nu} \equiv 0$. The equation $\bar{R}^\lambda_{\mu\sigma\nu} \equiv 0$ becomes our theory's first and main field equation.

5. Fifth – because the affine connections are now tensors, it is possible to construct for ourself – by way of choice (gauge constraint) – an effective geometry which is such that the non-linear terms $\bar{\Gamma}^\delta_{\mu\nu}\bar{\Gamma}^\lambda_{\delta\alpha}$ and $\bar{\Gamma}^\lambda_{\delta\nu}\bar{\Gamma}^\delta_{\mu\alpha}$ in the curvature tensor $\bar{R}^\lambda_{\mu\sigma\nu}$ vanish identically. This gauge choice results in three separate linear curvature tensors making up the resultant curvature tensor, namely $R^\lambda_{\mu\sigma\nu}$,

$\mathcal{T}^\lambda_{\mu\sigma\nu}$, and $Q^\lambda_{\mu\sigma\nu}$.

6. Sixth – the main field equation $\bar{R}^\lambda_{\mu\sigma\nu} \equiv 0$ is split into parts as $R^\lambda_{\mu\sigma\nu} = \mathcal{T}^\lambda_{\mu\sigma\nu}$ where $R^\lambda_{\mu\sigma\nu}$ is the Riemann curvature tensor and $\mathcal{T}^\lambda_{\mu\sigma\nu}$ the material curvature tensor.

7. Seventh – a set of gauge conditions (constraints) are then deliberately introduced – i.e. conditions which, when used in conjunction with the source-coupled field equation $R^\lambda_{\mu\sigma\nu} = \mathcal{T}^\lambda_{\mu\sigma\nu}$, yield for us the desired source-coupled Maxwell Geometrodymanic Equations [14].

8. Ante-penultimate – we split each of the curvature tensors $R^\lambda_{\mu\sigma\nu}$ and $\mathcal{T}^\lambda_{\mu\sigma\nu}$ into three parts each of which are also tensors.

9. Penultimate – we deduce the resultant field equations by relating each of the three tensors making up the Riemann curvature tensor $R^\lambda_{\mu\sigma\nu}$ to the three parts making up the material curvature tensor $\mathcal{T}^\lambda_{\mu\sigma\nu}$, wherefrom we obtain the first and second Maxwell's field equations [14], albeit in the realm of gravitomagnetism.

The above nine steps are an executive summary of the road leading to the theory here laid down. There is not much to say any further regarding the construction and architecture of the theory, except that we have given gravitomagnetism a fundamental geometric justification that we hope will lead researchers to reconsider gravitomagnetism as a fundamental phenomenon to be considered separately and independently as a physical phenomenon.

8.2 Gauge conditions

In total, the theory has required five gauge conditions for its architecture and design. These gauge conditions are presented in (37), (44), (79a), (79b), and (79c). Of these gauge conditions, (44) is the only natural gauge condition, while the rest are gauge constraints. The solo natural gauge is necessary in order that on a global level, the theory meets our most sought for requirement – of a geometry whose vectors during parallel transport in spacetime will have both their lengths and angles remain invariant. The gauge constraints (37), (79a), (79b), and (79c) have been instituted (imposed) so that we obtain a theory whose resulting equations have the form that we desire or that we are used to – which in this case, is the Maxwell form [14].

8.3 Quantum theory of gravity

Lastly, as our final word, we will briefly touch on the long sought – albeit elusive and contentious – dream of attaining a *Quantum Theory of Gravity* (QTG). Given the obvious similarities not only in the formulae of Sir Isaac Newton's universal law of gravitation $F_g = -G\mathcal{M}_g m_g / r^2$ and Coulomb's electrostatic law $F_e = Q q / 4\pi\epsilon r^2$, but in the two physical phenomenon themselves, we can learn one or two things from QED if we are to one day find a quantum mechanical description of the gravitational field.

For example, if we are to accept the thesis presented herein – this would mean that, like electricity, gravity is represented by a four-vector field. From this deduction, logically and intuitively, it would appear that the same method(s) used to quantize the electrodynamic phenomenon – can (and must) be applied somehow to the much sought for quantization program of the gravitational field. We know very well that QED is built on the fundamental soils of three very beautiful equations, namely the Dirac equation [63, 64] and Maxwell's two equations of electrodynamics [14], i.e.:

$$
\begin{aligned}
\imath\hbar\,\gamma_\mu\partial^\mu\psi &= m_0 c\psi & \text{(a)} \\
\partial^\mu F_{\mu\nu} &= \mu_0 J_\nu & \text{(b)} \\
\partial_\lambda F_{\mu\nu} + \partial_\nu F_{\lambda\mu} + \partial_\mu F_{\nu\lambda} &= 0 & \text{(c)}
\end{aligned}
\tag{86}
$$

where (86a) is the Dirac equation [63, 64] and (86b & c) are Maxwell's two equations of electrodynamics [14] respectively. In the Dirac equation (86a), γ_μ, m_0, and ψ are the usual four 4×4 Dirac matrices, the rest mass of the particle, and the four-component Dirac wavefunction, respectively.

Thus, in much the same manner, the gravitational field might be quantizable *via* the quantization of the gravitational four-vector field $\mathcal{A}_\mu$, in much the same way the electromagnetic four-vector A_μ has been quantized in QED under the scheme of the three equations given in (86). In order for this, the Dirac equation will have to be replaced by its curved spacetime equivalent. In [65], we did propose such a curved spacetime version of the Dirac equation, namely $\imath\hbar\gamma_\mu^{(a)}\mathcal{A}_\mu\partial^\mu\psi = m_0 c\psi$, and in our search for a QTG, we shall take this equation as the appropriate curved spacetime Dirac equation. Thus, we propose that the three equations to be used in the quantization program are:

$$
\begin{aligned}
\imath\hbar\,\gamma_\mu^{(a)}\mathcal{A}_\mu\partial^\mu\psi &= m_0 c\psi & \text{(a)} \\
\partial^\mu\mathfrak{F}_{\mu\nu} &= -\widetilde{\mu}\mathfrak{J}_\nu & \text{(b)} \\
\partial_\nu\mathfrak{F}_{\mu\sigma} + \partial_\sigma\mathfrak{F}_{\nu\mu} + \partial_\mu\mathfrak{F}_{\sigma\nu} &\equiv 0\,. & \text{(c)}
\end{aligned}
\tag{87}
$$

At the time when the curved spacetime Dirac equation (87a) was proposed, we where not sure how to identify the gravitational four-vector field $\mathcal{A}_\mu$ because we had not conceived of the gravitational field as capable of being described by a four-vector. But after the fundamental work of Behera [32] and Heras [43], and what we have presented herein, we are more than convinced that the gravitational field must submit to a four-vector description as suggested herein and e.g. by Heras [43], Behera [32], Heaviside [29, 30] and Maxwell [14].

8.4 In closing

In closing, allow us to say that as already stated a number of times, the purpose of the present paper has been to show that gravitomagnetism can be given a geometric description $\overline{g}_{\mu\nu} = e^{2\phi} g_{\mu\nu}$ on spacetime in exactly the same manner as Einstein gave gravity a geometric description on spacetime *via*

the metric tensor $g_{\mu\nu}$. For fear of digression and loss of focus, we have avoided going deeper in the many areas that this paper can possibly touch. We shall be making follow-up work which will dwell on these matters. We are very much aware of these many areas and we have not even mentioned some of them but silently passed as though we are not aware of them – this has been done intentionally. Further, for the same reasons, we have not done a serious comparative analysis of the present ideas with similar attempts in the literature. We must say that, the present paper is already an unavoidably lengthy one, so mush so that there really is no need to burden you our reader with more material. This can efficiently be done in separate papers in the future.

Dedications

We dedicate this paper to all the *Weylians* out there who, in the pristine of their privacy, took some time out to observe the *First Centenary of Weyl Gravitation*. Additionally, we take this opportunity to pay a befitting homage and tribute to the great personage, pristine intellect and esoteric genius of *Herr Professorie Dr. Herman Klaus Hugo Weyl* (1885-1955).

References

1. Weyl H. K. H. Gravitation und Elektrizität, Sitzungsber. *Preuss. Akad. Wiss*, 1918, v. 26, 465–478.

2. Einstein A. A Generalisation of the Relativistic Theory of Gravitation. *Ann. Math.*, 1945, v. 46, 578–584.

3. Einstein A and Straus E. G. *Ann. Math.*, 1946, v. 47, 731. see also: Einstein A. *Rev. Mod. Phys.*, 1948, v. 20, 35. Einstein A. *Can. J. Math.*, 1950, v. 2, 120.

4. Eddington A. S. The Mathematical Theory of Relativity. Cambridge, 1924.

5. Schrödinger E. R. J. A. The Final Affine Field Laws I. *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*, 1945, v. 51, 163–171.

6. Schrödinger E. R. J. A. The Final Affine Field Laws. II. *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*, 1945, v. 51, 205–216.

7. Schrödinger E. R. J. A. The Final Affine Field Laws. III. *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*, 1948, v. 52, 1–9.

8. Nyambuya G. G. Unified Field Theory. *Apeiron*, 2007, v. 14 (3), 320–361.

9. Nyambuya G. G. Unified Field Theory. *Apeiron*, 2008, v. 15 (1), 1–24.

10. Nyambuya G. G. Unified Field Theory in Nutshell – Ellicit Dreams of a Final Theory. *J. Mod. Phys.*, 2014, v. 5 (16), 1733–1766.

11. Straub W. O. On the Failure of Weyl's 1918 Theory. viXra: 1401.0168v4, 1–6.

12. Christoffel E. B. Ueber die Transformation der Homogenen Differentialausdrücke zweiten Grades. *Jour. für die reine und angewandte Mathematik B*, 1869, v. 70, 46–70.

13. Einstein A. Uber das Relativitätsprinzip und die aus demselben gezogenen Folgerungen. *Jahrbuch der Radioaktivität und Electronik*, 1907, v, 4, 411-462. Translation by Schwartz H. M. (1977), *Am. J. Phys.*, 1977, v. 45, 10.

14. Maxwell J. C. A Dynamical Theory of the Electromagnetic Field. *Phil. Trans. Royal Soc.*, 1865, v. 155, 459–512.

15. Adler R. J., Chen P., and Varani E. Gravitomagnetism and Spinor Quantum Mechanics. *Phys. Rev. D*, 2012, v. 85, 025016.

16. Exirifard Q. Gravito-Magnetic Force in Modified Newtonian Dynamics. *Journal of Cosmology and Astroparticle Physics*, 2013, v. 2013 (08), 046–046.

17. Mashhoon B. Gravitoelectromagnetism: A Brief Review. arXiv: gr-qc/0311030v2.

18. Nordtvedt K. Gravitomagnetic Interaction and Laser Ranging to Earth Satellites. *Phys. Rev. Lett.*, 1988, v. 61, 2647–2649.

19. Vieira R. S. and Brentan H. B. Covariant Theory of Gravitation in the Framework of Special Relativity. *The European Physical Journal Plus*, 2018, v. 133 (4), 165–190.

20. Chicone C. and Mashhoon B. Gravitomagnetic Jets. *Phys. Rev. D*, 2011, v. 83, 064013.

21. Iorio L. Long-term Classical and General Relativistic Effects on the Radial Velocities of the Stars Orbiting Sgr A*. *MNRAS*, 2011, v. 411 (1), 453–463.

22. Iorio L., Lichtenegger H. I. M., Ruggiero M. L., and Corda C. Invited Review: Phenomenology of the Lense-Thirring Effect in the Solar System. *Astrophys. Space Sci.*, 2011, v. 331, 351–395.

23. Mashhoon B., Gronwald F., and Lichtenegger H. I. M. Gravitomagnetism and the Clock Effect. *Lecture Notes in Physics*, 2001, v. 562, 83.

24. Soffel M., Klioner S., Müller J., and Biskupek L. Gravitomagnetism and Lunar Laser Ranging. *Phys. Rev. D*, 2008, v. 78, 024033.

25. Einstein A. Grundlage der allgemeinen Relativitätstheorie (The Foundation of the General Theory of Relativity). *Ann. Phys. (Leipzig)*, 1916, v. 49 (7), 769–822.

26. Einstein A. Näherungsweise Integration der Feldgleichungen der Gravitation (Approximative Integration of the Field Equations of Gravitation). *Preussische Akademie der Wissenschaften, Sitzungsberichte (Part 1)*, 1916, 688–696.

27. Einstein A. Grundgedanken der allgemeinen Relativitätstheorie und Anwendung dieser Theorie in der Astronomie (Fundamental Ideas of the General Theory of Relativity and the Application of this Theory in Astronomy). *Preussische Akademie der Wissenschaften, Sitzungsberichte (Part 1)*, 1915, p. 315.

28. Einstein A. Zur allgemeinen Relativitätstheorie (On the General Theory of Relativity). *Preussische Akademie der Wissenschaften, Sitzungsberichte (Part 2)*, 1915, 778–786, 799–801.

29. Heaviside O. A Gravitational and Electromagnetic Analogy. *The Electrician*, 1893, v. 31, 281–282 & 359.

30. Heaviside O. Electromagnetic Theory. The Electrician Printing and Publishing Co., London, 1894, pp. 455–465.

31. Jefimenko O. D. Causality, Electromagnetic Induction and Gravitation: A Different Approach to the Theory of Electromagnetic and Gravitational Fields. Electret Scientific, Star City, 2000. ISBN 0-917406-22-2/ 0-917406-23-0.

32. Behera H. Newtonian Gravitomagnetism and Analysis of Earth Satellite Results. arXiv: gr-qc/0510003v2.

33. Nyambuya G. G. A Perdurable Defence to Weyl's Unified Theory. *J. Mod. Phys.*, 2014, v. 5 (14), 1244–1253.

34. Einstein A. Zur Elektrodynamik bewegter Körper. *Ann. der Phys.*, 1905, v. 17, 891.

35. Gödel K. An Example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation. *Reviews of Modern Physics*, 1949, v. 21 (3), 447–450.

36. Carroll S. M. Lecture Notes on General Relativity. arXiv: gr-qc/9712019v1.

37. Einstein A. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Preussische Akademie der Wissenschaften, Sitzungsberichte (Part 1)*, 1917, 142–152

38. Pais A. Subtle is the Lord. Oxford Univ. Press, 2005.

39. Einstein A and Grommer J. Allgemeine Relativitätstheorie und Bewegungsgesetz (General Theory of Relativity and the Law of Motion). *Sitzungsberichte der Preussischen Akademie der Wissenschaften (Part 1)*, 1927, 2–13, 235–245.

40. Berry M. Paul Dirac: The Purest Soul in Physics. *Physics World*, 1998, v. 11 (2), 36–40.

41. Kragh H. S. Dirac. Cambridge University Press, 2003. ISBN-13/ISBN-10: 978-0521380898/0521380898.

42. Mehra J. The Genius and Creativity of a Pure Soul. *Physics World*, 1991, v. 4 (4), 51–52.

43. Heras J. A. Can Maxwell's Equations Be Obtained from the Continuity Equation? *Am. J. Phys.*, 2007, v. 75, 652.

44. Weyl H. K. H. Elektron und Gravitation I. *Z. Phys.*, 1927, v. 56, 330–352.

45. Weyl H. K. H. Gravitation and the Electron. *Proc. Nat. Acad. Sci.*, 1927, v. 15, 323–334.

46. Das A., Kaku M., and Townsend P. K. Unified Approach to Matter Coupling in Weyl and Einstein Supergravity. *Phys. Rev. Lett.*, 1978, v. 40, 1215–1218.

47. Dirac P. A. M. Long Range Forces and Broken Symmetries. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 1973, v. 333 (1595), 403–418.

48. Gregorash D. and Papini G. Weyl-Dirac Theory with Torsion. *Il Nuovo Cimento B (1971-1996)*, 1980, v. 55 (1), 37–51.

49. Hochberg D. and Plunien G. Theory of Matter in Weyl Spacetime. *Phys. Rev. D*, 1991, v. 43, 3358–3367.

50. Maki T., Kan N., and Shiraishi K. Dirac-Born-Infeld-Einstein Theory with Weyl Invariance. *J. Mod. Phys.*, 2012, v. 3, 1081–1087.

51. Pricop M., Răut M., Borsos Z., Baciu A., and Agop M. Holographic-Type Gravitation *via* Non-Differentiability in Weyl-Dirac Theory. *J. Mod. Phys.*, 2013, v. 4, 165–171.

52. Romero C., Fonseca-Neto J. B., and Pucheu M. L. General Relativity and Weyl Geometry. *Class. Quantum Grav.*, 2012, v. 29 (15), 155015.

53. Wheeler J. T. Weyl Gravity as General Relativity. *Phys. Rev. D*, 2014, v. 90, 025027.

54. Yuan F. F. and Huang Y. C. A Modified Variational Principle for Gravity in the Modified Weyl Geometry. *Class. Quantum Grav.*, 2013, v. 30 (19), 195008.

55. Einstein A. Die Feldgleichungun der Gravitation. *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin (Part 2)*, 1915, 844–847.

56. Gönner H. F. M. On the History of Unified Field Theories. *Living Rev. Relativ.*, 2004, v. 7 (2), 1–153.

57. Gönner H. F. M. On the History of Unified Field Theories. Part II. (ca. 1930 – ca. 1965). *Living Rev. Relativ.*, 2014, v. 17 (5), 1–241.

58. Weinberg S. The Quantum Theory of Fields, Vol. 1. Cambridge University Press, Cambridge, UK, 2005, p. 225.

59. Zee A. Quantum Field Theory in a Nutshell. Princeton Univ. Press, 2010, p. 98.

60. Bianchi L. Sui Simboli a Quattro Indici e Sulla Curvatura di Riemann. *Rend. Acc. Naz. Lincei*, 1902, v. 11 (5), 3–7.

61. Voss A. Zur Theorie der Transformation Quadratischer Differentialausdrücke und der Krümmung Höherer Mannigfaltigkeiten. *Mathematische Annalen*, 1880, v. 16 (2), 129–179.

62. Walter I. Einstein. Simon + Schuster Inc., 2008. ISBN: 978-1-8473-9589-4 (eBook)/978-1-84739-054-7.

63. Dirac P. A. M. The Quantum Theory of the Electron. *Proc. Roy. Soc. (London)*, 1928, v. A117, 610–612.

64. Dirac P. A. M. The Quantum Theory of the Electron II. *Proc. Roy. Soc. (London)*, 1928, v. A118, 351–361.

65. Nyambuya G. G. New Curved Spacetime Dirac Equations. *Found. Phys.*, 2008, v. 37 (7), 665–677.

Can Nano-Materials Push Off the Vacuum?

M. E. McCulloch

Plymouth University, UK. E-mail: mike.mcculloch@plymouth.ac.uk

The theory of quantised inertia (QI), which predicts galaxy rotation without dark matter, also predicts that electromagnetic energy input into an asymmetric cavity perceives a gradient in the quantum vacuum in the cavity producing a force on that cavity. Here it is shown that if the cavity is less than 129 nm in scale, then no input power is needed and the predicted thrust can be comparable to gravity. Arrays of these nano-cavities could produce a self-thrusting material.

1 Introduction

Many astrophysical observations show that stars at the outer edges of galaxies orbit far too fast to be gravitationally bound to the galaxy [1, 2] and an identical phenomenon is observed for globular clusters [3] and wide binaries [4]. On a much smaller scale, some laboratory experiments have shown that asymmetric metal cavities of various types with strong electromagnetic fields resonating within them (emdrives) show an unexpected thrust towards their narrower ends [5, 6].

All these phenomena can be predicted by a theory called quantised inertia, which assumes that the inertial force arises because the Rindler horizon that objects see when they accelerate damps the excited zero point field (Unruh radiation) behind them creating an imbalance which pushes them back against their original acceleration [7, 8]. This model successfully predicts galaxy and wide binary rotations without any adjustment [9, 10]. Quantised inertia also predicts that an artificial horizon can be produced when high acceleration matter or electromagnetic radiation is confined inside an asymmetric cavity, producing a new kind of thrust [11, 12] that may already have been seen in the emdrive. It was pointed out by [13] that using light and supermirrors to contain it, might enhance this force.

It is shown here that QI also predicts that if the asymmetric metal cavities are as small as 129 nm then a thrust comparable to gravity can be obtained even from the unexcited zero point field. This implies that if a material was constructed with arrays of asymmetric nano-cavities, then the force would be enough to levitate that material.

2 Method & result

We start with Heisenberg's uncertainty principle for a single photon inside a double-cavity that has a wide part and a narrow part (see Figure 1). A photon oscillates repeatedly along a distance d between the wide and narrow cavities as shown by the arrow. The uncertainty principle states that the uncertainty in momentum (Δp) and position (Δx) of the photon in each cavity is

$$\Delta p \Delta x \geq \hbar/2 \,. \tag{1}$$

The uncertainty in position is assumed, in quantised inertia, to be the size of the cavity the photon is in. [14,15] pointed
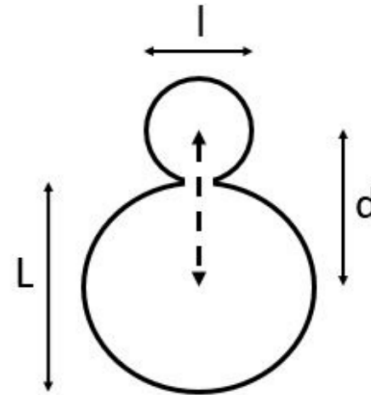

Fig. 1: The asymmetrical metal cavity. A photon moves back and forth along the dashed arrow.

out that Heisenberg's original form for the uncertainty principle intended an equal sign, not an inequality so that in the wide cavity we can write

$$\Delta p_w = \frac{\hbar}{2L} \tag{2}$$

and for the narrow cavity

$$\Delta p_n = \frac{\hbar}{2l} \,. \tag{3}$$

The force is the change of momentum with time

$$F = \frac{\Delta p}{\Delta t} = \frac{c(\Delta p_n - \Delta p_w)}{d} = \frac{\hbar c}{2\left(\frac{L}{2} + \frac{l}{2}\right)}\left(\frac{1}{l} - \frac{1}{L}\right). \tag{4}$$

If we assume that the width of the smaller cavity is half that of the larger ($l = L/2$) then

$$F = ma = \frac{2\hbar c}{3L^2} \,. \tag{5}$$

The mass of the cavity, assuming it is two hollow spheres, is $m = 5\pi L^2 \rho \delta / 4$ where ρ is the density of the metal walls and δ is their thickness. So

$$L^2 = \frac{2\hbar c}{3ma} = \frac{8\hbar c}{15\pi L^2 \rho \delta a} \,. \tag{6}$$

Rearranging, we can now calculate the size of cavity at which the energy solely from the zero point field ($\hbar$) is enough to produce acceleration a

$$L = \sqrt[4]{\frac{8\hbar c}{15\pi\rho\delta a}}. \qquad (7)$$

Assuming that the density of the metal is 2000 kg/m^3, its thickness is 1 mm and the acceleration to be overcome is that of gravity at the Earth's surface, $g = 9.8$ m/s^2, then we get

$$L = 129\,\text{nm}. \qquad (8)$$

The implication is that if we build an asymmetric metal cavity such as that shown in Figure 1, with its narrow end upwards and on a scale of 129 nm or less, then it should levitate simply from the already-present zero point field without any input power.

3 Discussion

It follows from the above that if a material can be manufactured that is composed of an array of asymmetric nano-structures of size 129 nm or less then the material will levitate without input power.

One difficulty will be that, on the nanoscales considered here, other thermal or plasmonic effects will become important so the effectiveness of this approach will be dependent on these other effects cancelling out.

4 Conclusions

Quantised inertia predicts that asymmetric metal cavities make a gradient in the quantum vacuum, causing thrust.

The smaller the cavity, the larger the predicted thrust. At scales of 129 nm, the thrust equals gravity at the Earth's surface.

If a material can be constructed with arrays of such asymmetric nano-cavities then it should levitate without input power.

Acknowledgements

References

1. Zwicky F. On the masses of nebulae and clusters of nebulae. *The Astrophysical Journal*, 1937, v. 86, 217.

2. Rubin V.C., Ford W.K. *The Astrophysical Journal*, 1970, v. 159, 379.

3. Scarpa R., Marconi G., Gimuzzi R. and Carraro G. *A& A*, 2007, v. 462, L9.

4. Hernandez X., Cortes R.A.M., Allen C. and Scarpa R. Challenging a Newtonian prediction through Gaia wide binaries. *IJMP-D*, 2018, v. 28, 8, 1950101.

5. Shawyer R. Microwave propulsion — progress in the emdrive programme. *59th International Astronautical Conference*, IAC-2008, Glasgow, UK, 2008.

6. White H., March P., Lawrence J., Vera J., Sylvester A., Brady D. and Bailey P. *J. Propulsion and Power*, 17 Nov 2016.

7. McCulloch M.E. Modelling the Pioneer anomaly as modified inertia. *Mon. Not. Roy. Astron. Soc.*, 2007, v. 376, 338–342.

8. McCulloch M.E. Inertia from an asymmetrical Casimir effect. *EPL*, 2013, v. 101, 59001.

9. McCulloch M.E. Testing quantised inertia on galactic scales. *Astrophys. Space Sci.*, 2012, v. 342, no. 2, 575–578.

10. McCulloch M.E. and Lucio J. Testing Newton/GR, MoND & quantised inertia on wide binaries. *Astrophys. Space Sci.*, 2019, v. 364, 121.

11. McCulloch M.E. Testing quantised inertia on the emdrive, *EPL*, 2015, v. 111, 60005; arXiv.org/abs/1604.03449.

12. McCulloch M.E. Propellant-less propulsion from quantised inertia. *J. Space Explor.*, 2018, v. 7(3).

13. Taylor T. Propulsive forces using high-Q asymmetric high energy laser resonators. *JBIS*, 2017, v. 70, 238–243.

14. Süssmann G. Uncertainty relation: from inequality to equality. *Z. Naturforsch.*, 1997, v. 52a, 49–52.

15. Hall M.J.W. Exact uncertainty relations: physical significance. *Phys. Rev. A*, 2001, v. 64, 052103–1–10.

————————————

New Approach to Measurement in Quantum Tomography

Artur Czerwinski

Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Grudziadzka 5, 87-100 Torun, Poland.
E-mail: aczerwin@umk.pl

In this article we propose a new approach to quantum measurement in reference to the stroboscopic tomography. Generally, in the stroboscopic approach it is assumed that the information about the quantum system is encoded in the mean values of certain hermitian operators $Q_1, ..., Q_r$ and each of them can be measured more than once. The main goal of the stroboscopic tomography is to determine under which conditions one is able to reconstruct the initial density matrix $\rho(0)$ on the basis of the measurement results $\langle Q_i \rangle_{t_j}$. In this paper we propose to treat every complex matrix as a measurable operator. This generalized approach to quantum measurement may bring some improvement into the models of stroboscopic tomography.

1 Introduction

In this paper by $\mathcal{H}$ we shall denote the Hilbert space and we shall assume that $dim\mathcal{H} = n < \infty$. By $B(\mathcal{H})$ we shall denote the complex vector space of all bounded linear operators in $\mathcal{H}$. The space $B(\mathcal{H})$ is isomorphic with the space of all complex matrices that shall be represented by $\mathbb{M}_n(\mathbb{C})$. Finally, $B_*(\mathcal{H})$ shall refer to the real vector space of all hermitian (self-adjoint) operators on $\mathcal{H}$. The elements of $B_*(\mathcal{H})$ shall be called observables.

The term *quantum state tomography* refers to a wide variety of methods and approaches which aim to reconstruct the accurate representation of a quantum system by performing a series a measurements. Among many different approaches to quantum tomography, one can especially mention the so-called static model of tomography, which requires $n^2 - 1$ measurements of different observables taken at time instant $t = 0$ (see more in [1–3]). A paper published in 2011 initiated another approach to quantum tomography which is based on weak measurement. The paper revealed that the wave function of a pure state can be measured in a direct way [4]. Further papers proved that this approach can be generalized also for mixed state identification [5].

In this paper we follow yet another approach to quantum tomography – the so-called stroboscopic tomography which originated in 1983 in the article [6]. Subsequently, the approach was developed in other papers, such as [7–9]. The assumption that lies at the very foundation of this method claims that the evolution of an open quantum system can be expressed by a master equation of the form

$$\dot{\rho}(t) = \mathbb{L}[\rho(t)], \tag{1}$$

where the operator $\mathbb{L}$ is called the generator of evolution and its most general form have been introduced in [10]. In order to determine the initial density matrix $\rho(0)$ one assumes to have a set of identically prepared quantum systems which evolve according to the master equation with the generator $\mathbb{L}$. Each system can be measured only once, because any measurement, generally, influences the state.

The other underlying assumption connected with the stroboscopic approach is that the knowledge about the quantum system is provided by mean values of certain observables $\{Q_1, ..., Q_r\}$ (obviously $Q_i^* = Q_i$) such that $r < n^2 - 1$. These mean values are mathematically expressed as

$$\langle Q_i \rangle_t = \text{Tr}(Q_i \, \rho(t)) \tag{2}$$

and are assumed to be achievable from an experiment. If we additionally assume that the knowledge about the evolution enables us to perform measurements at different time instants $t_1, ..., t_g$, we get from an experiment a matrix of data $[\langle Q_i \rangle_{t_j}]$, where $i = 1, ..., r$ and $j = 1, ..., g$. The fundamental question of the stroboscopic tomography that one asks is: whether the matrix of experimental data is sufficient to reconstruct the initial density matrix $\rho(0)$? Other problems relate to the minimal number of observables and time instants, the properties of the observables and the choice of time instants. In general the conditions under which it is possible to reconstruct the initial state have been determined and can be found in [6–8].

Compared with the static model of tomography, the stroboscopic approach makes it possible to decrease significantly the number of different observables that are necessary to perform quantum state tomography. From the experimental point of view, it means that in the static model one needs to prepare $n^2 - 1$ different experimental systems (e.g. for $dim \, \mathcal{H} = 4$ one would need to measure 15 different quantities), which seems rather unrealistic. Therefore, the stroboscopic approach appears to have an advantage over the static model as it aims to reduce the number of distinct observables.

2 Generalized observables and measurement results

According to one of the most fundamental concepts of quantum mechanics, to every physical quantity we can assign a hermitian operator which is called an observable. Thus, when talking about measurements in the context of the stroboscopic tomography, we consider mean values of certain hermitian operators [6]. In general, any hermitian operator can be de-

composed according to the spectral theorem:

$$Q = \sum_i \lambda_i P_i. \tag{3}$$

where P_i is the projector onto the eigenspace of Q with the eigenvalue λ_i [11]. Physically speaking, the possible results of measurement correspond to the eigenvalues of Q, whereas the probability of getting the result λ_i (upon measuring the state ρ) can be expressed as:

$$p_i = \mathrm{Tr}(P_i \rho). \tag{4}$$

Finally, we can compute the expectation value of Q as:

$$\langle Q \rangle = \sum_i \lambda_i \, p_i = \mathrm{Tr}(Q \rho), \tag{5}$$

which gives the famous formula for the mean value of any observable.

In other words, any observable is associated with a projective measurement, which stems from the spectral theorem. The main goal of this section is to prove that this approach to measurement can be generalized in such a way that any complex matrix $A \in \mathbb{M}_n(\mathbb{C})$ can be considered a measurable operator.

We shall formulate and employ the following theorem.

Theorem 1. *(Hermitian decomposition of a complex matrix) For any matrix $A \in \mathbb{M}_n(\mathbb{C})$ there exist two matrices $Q, R \in B_*(\mathcal{H})$ such that the matrix A can be decomposed as*

$$A = Q + i R. \tag{6}$$

Proof. Let us first denote $A = [a_{ij}]$ and since in general $a_{ij} \in \mathbb{C}$ we can put

$$a_{ij} = \mathrm{Re}\, a_{ij} + i \,\mathrm{Im}\, a_{ij}. \tag{7}$$

Moreover we can denote $Q = [q_{ij}]$ and $R = [r_{ij}]$. Then we shall define the entries of the matrices Q and R in the way:

$$q_{ij} := \frac{\mathrm{Re}\, a_{ij} + \mathrm{Re}\, a_{ji}}{2} + i\, \frac{\mathrm{Im}\, a_{ij} - \mathrm{Im}\, a_{ji}}{2}, \tag{8}$$

$$r_{ij} := \frac{\mathrm{Im}\, a_{ij} + \mathrm{Im}\, a_{ji}}{2} + i\, \frac{\mathrm{Re}\, a_{ji} - \mathrm{Re}\, a_{ij}}{2}. \tag{9}$$

One can easily notice that $\overline{q_{ij}} = q_{ji}$ and $\overline{r_{ij}} = r_{ji}$. Therefore $Q, R \in B_*(\mathcal{H})$.

Furthermore, one can check that

$$q_{ij} + i r_{ij} = \frac{\mathrm{Re}\, a_{ij} + \mathrm{Re}\, a_{ji}}{2} + i\, \frac{\mathrm{Im}\, a_{ij} - \mathrm{Im}\, a_{ji}}{2} + \\ + i\, \frac{\mathrm{Im}\, a_{ij} + \mathrm{Im}\, a_{ji}}{2} + \frac{\mathrm{Re}\, a_{ij} - \mathrm{Re}\, a_{ji}}{2} = a_{ij}, \tag{10}$$

which implies that

$$A = Q + i R. \tag{11}$$

$\square$

The above theorem states that every complex matrix $A \in \mathbb{M}_n(\mathbb{C})$ can be uniquely decomposed into two hermitian matrices. In other words, every complex matrix can be regarded as a pair of observables (hermitian matrices), i.e.

$$A \rightarrow (Q_1, Q_2), \text{ where } Q_1, Q_2 \in B_*(\mathcal{H}). \tag{12}$$

Since in general any observable is considered measurable, therefore, any complex matrix can also be considered a measurable operator.

In this paper it has been proven that for any $A \in \mathbb{M}_n(\mathbb{C})$ there exist two observables $Q_1, Q_2 \in B_*(\mathcal{H})$ such that

$$A = Q_1 + i\, Q_2. \tag{13}$$

If we generalize the idea of quantum measurement, we can define the mean value of any operator $A \in \mathbb{M}_n$ measured upon a quantum system characterized by a density matrix $\rho(t)$. Such a quantity, denoted by $\langle A \rangle_t$, shall be defined in the following way:

$$\langle A \rangle_t := \mathrm{Tr}[A\rho(t)] = \mathrm{Tr}\left[(Q_1 + i\, Q_2)\rho(t)\right]. \tag{14}$$

Taking into account the fact that *trace* is linear, one obtains

$$\langle A \rangle_t = \mathrm{Tr}[Q_1\rho(t)] + i\, \mathrm{Tr}[Q_2\rho(t)], \tag{15}$$

which can be equivalently presented as

$$\langle A \rangle_t = \langle Q_1 \rangle_t + i\, \langle Q_2 \rangle_t. \tag{16}$$

One can observe that if we generalize the idea of quantum measurement in such a way that we treat any complex matrix $A \in \mathbb{M}_n(\mathbb{C})$ as a measurable operator, the mean value of A is a complex number such that its real and imaginary parts are mean values of the observables Q_1, Q_2 which appear in the hermitian decomposition of A. Therefore, the measurement of any complex operator A can be understood as the measurement of two physical quantities that are mathematically represented by the hermitian matrices Q_1, Q_2.

3 Connection with the stroboscopic tomography

When considering problems in the stroboscopic tomography, one needs to bear in mind the necessary condition that the set of observables $Q_1, Q_2, ..., Q_r$ has to satisfy so that an open quantum system with dynamics given by (1) will be reconstructible.

Theorem 2. *An open quantum system with evolution given by Eq. 1 is $(Q_1, ...Q_r)$-reconstructible if and only if the operators Q_i satisfy the condition [6,7]*

$$\bigoplus_{i=0}^{r} K_\mu(\mathbb{L}, Q_i) = B_*(\mathcal{H}), \tag{17}$$

where $\bigoplus$ refers to the Minkowski sum, μ is the degree of the minimal polynomial of $\mathbb{L}$ and $K_\mu(\mathbb{L}, Q_i)$ denotes the Krylov subspace which standard definition reads:

$$K_\mu(\mathbb{L}, Q_i) := \mathrm{Span}\{Q_i, \mathbb{L}^* Q_i, (\mathbb{L}^*)^2 Q_i, ..., (\mathbb{L}^*)^{\mu-1} Q_i\}. \tag{18}$$

In reference to this condition for observability of a quantum system we can propose the following theorem.

Theorem 3. *Assume that the set of hermitian matrices denoted by $\{\lambda_1, \lambda_2, ..., \lambda_{n^2}\}$ constitutes a basis in the space of all hermitian operators $B_*(\mathcal{H})$, where $n = dim\,\mathcal{H}$. Then they also constitute a basis in the space of all linear operators $\mathbb{M}_n(\mathbb{C})$.*

Proof. Taking into account the assumption, one can write:

$$\forall_{Q \in B_*(\mathcal{H})} \quad \exists_{\alpha_1,...,\alpha_{n^2} \in \mathbb{R}} \quad Q = \sum_{k=1}^{n^2} \alpha_k \lambda_k. \tag{19}$$

Then from the theorem on hermitian decomposition of a complex matrix it follows that $\forall_{A \in \mathbb{M}_n(\mathbb{C})} \ \exists_{Q,R \in B_*(\mathcal{H})}$ such that the matrix A can be decomposed as

$$A = Q + i\,R. \tag{20}$$

Assuming that Q has such decomposition as in (19) and taking R in the analogous form:

$$R = \sum_{k=1}^{n^2} \beta_k \lambda_k, \quad \beta_k \in \mathbb{R}, \tag{21}$$

matrix A can be represented as

$$A = \sum_{k=1}^{n^2} \alpha_k \lambda_k + i\left(\sum_{k=1}^{n^2} \beta_k \lambda_k\right), \tag{22}$$

which can be transformed into the form

$$A = \sum_{k=1}^{n^2} (\alpha_k + i\beta_l)\,\lambda_k. \tag{23}$$

Finally, the matrix A can be decomposed as

$$A = \sum_{k=1}^{n^2} z_k \lambda_k, \tag{24}$$

where $z_k \in \mathbb{C}$ and $z_k = \alpha_k + i\beta_k$.
From (24) one can easily draw the conclusion that the set of matrices $\{\lambda_1, \lambda_2, ..., \lambda_{n^2}\}$ is a basis in $\mathbb{M}_N(\mathbb{C})$. $\square$

The link between the above theorem and the stroboscopic tomography is that in (17), which expresses the necessary condition for observability, on the right hand side you can put either $B_*(\mathcal{H})$ or $B(\mathcal{H})$. On the basis of theorem 3 one can conclude that if certain operators span one of these spaces, they also have to span the other.

4 Summary

In this paper it has been proved that any complex matrix $A \in \mathbb{M}_n(\mathbb{C})$ can be uniquely determined by two hermitian matrices

(i.e. observables). In general, mean values of hermitian matrices can be obtained from an experiment (based on projective measurement). Thus, from this observation one can conclude that any complex matrix can be regarded as a measurable operator. The measurement of a complex matrix should be understood as the measurement of the mean values of two observables which determine the complex operator. The measurement result of a complex matrix is then a complex number which real and imaginary parts are obtained from an experiment. Further research is planned to investigate whether the generalized approach to measurable operators can improve the models of the stroboscopic tomography.

References

1. Altepeter J. B., James D. F. V., Kwiat P. G. 4 Qubit Quantum State Tomography. In: Paris M. G. A., Rehacek J., eds. Quantum State Estimation. Springer, Berlin, 2004, 111–145.

2. Alicki R., Lendi K. Quantum Dynamical Semigroups and Applications. Springer, Berlin, 1987.

3. Kimura G. The Bloch vector for N-level systems. *Phys. Lett. A*, 2003, v. 314 (5–6), 339–349.

4. Lundeen J. S., Sutherland B., Patel A., Stewart C., Bamber C. Direct measurement of the quantum wavefunction. *Nature*, 2011, v. 474 (7350), 188–191.

5. Wu S. State tomography via weak masurements. *Scientific reports*, 2013, v. 3, 1193.

6. Jamiolkowski A. The minimal Number of Operators for Observability of N-level Quantum Systems. *Int. J. Theor. Phys.*, 1983, v. 22 (4), 369–376.

7. Jamiolkowski A. On complete and incomplete sets of observables, the principle of maximum entropy – revisited. *Rep. Math. Phys.*, 2000, v. 46 (3), 469–482.

8. Jamiolkowski A. On a Stroboscopic Approach to Quantum Tomography of Qudits Governed by Gaussian Semigroups. *Open Syst. Inf. Dyn.*, 2004, v. 11 (1), 63–70.

9. Jamiolkowski A. Fusion Frames and Dynamics of Open Quantum Systems. In: Lyagushyn S., ed. Quantum Optics and Laser Experiments. InTech, Rijeka, 2012, 67–84.

10. Gorini V., Kossakowski A., Sudarshan E. C. G. Completely positive dynamical semigroups of n-level systems. *J. Math. Phys.*, 1976, v. 17 (5), 821–825.

11. Nielsen M. A., Chuang I. L. Quantum Computation and Quantum Information. Cambridge University Press, Cambridge, 2000.

The Physics of Lithospheric Slip Displacements in Plate Tectonics

Pierre A. Millette

E-mail: pierre.millette@uottawa.ca, Ottawa, Canada.

In this paper, we present physical calculations to support a mechanism of slip displacements of the lithosphere in the plate tectonics model of the earth sciences. In particular, for a lithospheric slip displacement to occur, a force must be applied to the lithospheric plate to overcome the force of static friction that is holding it in place on top of the asthenosphere. The magnitude of the required applied force can be generated by asteroid impact and is found to depend on the mass of the plate, the mass, velocity and angle of incidence of the asteroid, and the duration of the momentum transfer. The distance that is covered by the plate as a result of the lithospheric slip displacement is calculated and provides an explanation for observed sudden changes in direction and/or speed of plate motions. The model calculations presented in this paper provide a framework to analyze lithospheric slip displacements in plate tectonics resulting from asteroid impacts.

1 Introduction

In this paper, we present physical calculations to support a mechanism of slip displacements of the lithosphere in the plate tectonics model of the earth sciences [1–3]. The lithosphere consists of the Earth's crust of thickness ~10 km and the upper part of the mantle composed of rigid rocks of average density $\rho \sim 3.3\,\mathrm{gm/cm^3}$, with overall average thickness ~100 km [3] [4, p. 76], divided into the tectonic plates covering the surface of the Earth. It rests on the upper part of the asthenosphere of average density $\rho \sim 3.1\,\mathrm{gm/cm^3}$ [4, p. 70], which is plastic and subject to viscous flows due to the nature of the rocks and the heat and densities involved. The asthenosphere becomes more rigid and stronger with increasing depth in the mantle, with average density $\rho \sim 3.4 - 4.4\,\mathrm{gm/cm^3}$. The earth's crust is differentiated from the lithospheric part of the mantle by the Mohorovičić, usually referred to as the Moho, discontinuity. See Fig. 1.

Given the structure of the lithosphere and the plastic and viscous nature of the upper part of the asthenosphere (low-viscosity zone LVZ [5, pp. 11,181]), it is quite conceivable that the lithosphere could move over the asthenosphere by a slip displacement movement, given the appropriate applied force to initiate the process. We calculate the applied force that would be required to initiate this process, and the type and nature of displacement movements that could be generated by such an applied force.

2 Lithospheric slip displacements

As currently understood, plate tectonics is a convective process, thermally driven by colder lithospheric slabs sinking into the interior of the hotter mantle at subduction zones [4, p. 11]. Continental drift and plate tectonics are considered to be sufficient proof of convection in the upper mantle [6, pp. 207–211].

However, as pointed out by Price [4, p. 63], "the models ... are completely unusable to explain the abrupt changes of rate and direction of plate motion which are, from time to

Fig. 1: Cross-section of the layers of the Earth's upper mantle and crust (not to scale). Force model for the lithospheric slip displacement.
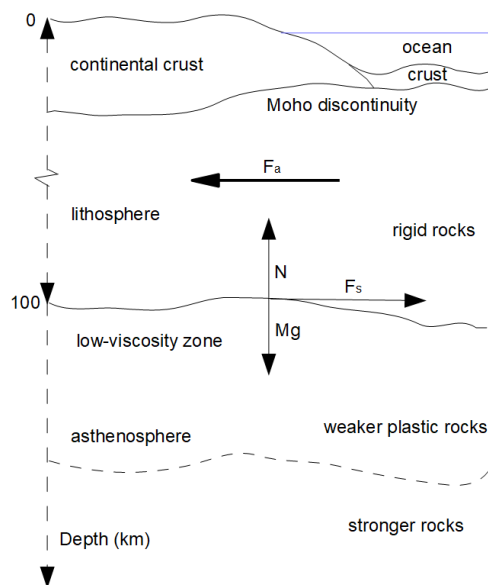
time, exhibited in the geological record". As stated in [7] quoted in Price [4, p. 191], "Unfortunately, we cannot reproduce the toroidal/poloidal partitioning ratios observed from the Cenozoic, *nor do our models explain apparently sudden plate motion changes that define stage boundaries*." [emphasis in Price]. A process of lithospheric slip displacement is needed to explain such sudden plate motions.

2.1 The force model

In this and subsequent sections, we seek to understand the lithospheric slip displacement process by performing order-of-magnitude simplified calculations. This first portion is a

simple force model (see Fig. 1).

We consider a tectonic plate of mass M resting on the asthenosphere with a static coefficient of friction μ_s. The force of static friction between the plate and the asthenosphere is then given by $F_s = -\mu_s N$, where the normal force N is given by $N = Mg$ where g is the acceleration due to gravity. Combining these quantities, the force of static friction F_s is then given by

$$F_s = -\mu_s Mg \,. \tag{1}$$

For the lithospheric slip displacement to occur, a force F_a must be applied to the plate to overcome the force of static friction that is holding it in place. This applied force must be greater than the force of static friction $F_a > F_s$, and substituting from (1), we obtain the slip condition

$$F_a > \mu_s Mg \,. \tag{2}$$

We consider a sample calculation for the North American plate as an order-of-magnitude estimate of the forces involved. The area of the North American plate is given by 58.8×10^6 km^2 [4, p. 7]. For an average thickness ~100 km and an average density $\rho \sim 3.3$ gm/cm^3 (see section 1), the mass of the North American plate is given by $M = 1.8 \times 10^{22}$ kg. Using these values and an estimated static coefficient of friction of 0.28 (greasy nickel) [8], the slip condition (2) then becomes

$$F_a > 5 \times 10^{22} \, \text{N} \,, \tag{3}$$

where N is the Newton unit of force. This estimated applied force slip condition could be higher in the case of a higher static coefficient of friction, but it would likely not exceed a factor of two higher (i.e. $F_a > 10^{23}$ N). For example, the static coefficient of friction between concrete and silty clay is estimated at 0.30-0.35 in [9]. This applied force required for a lithospheric slip displacement to occur is very significant.

The applied force provides the impulse to set the plate in motion. Once the plate is set in motion, the only force that is applicable is the force of kinetic friction between the plate and the asthenosphere which is slowing down the plate's movement. This force is given by $F_k = -\mu_k N$, where the normal force N is again given by $N = Mg$. The kinetic coefficient of friction μ_k is smaller than the static coefficient of friction μ_s. Combining these quantities, the force of kinetic friction F_k is then given by

$$F_k = -\mu_k Mg \,, \tag{4}$$

which decelerates the plate at the rate $a = -\mu_k g$. For the example previously considered, using an estimated kinetic coefficient of friction of 0.12 (greasy nickel) [8], the deceleration is given by $a = -1.2$ m s^{-2}. The deceleration could be greater in the case of a higher kinetic coefficient of friction, but it would likely not exceed a factor of two higher (i.e. $a = -2.4$ m s^{-2}). For example, the sliding (kinetic) coefficient of friction between cement and wet clay is estimated at 0.2 in [8].

2.2 The asteroid impact model

As we have seen in (3), the applied force required for a lithospheric slip displacement to occur is very significant. This magnitude of force would only be available in a collision process, such as the impact of an asteroid or comet with the plate. We use the term asteroid impact in a generic fashion to represent both asteroid and comet impacts. Neville Price has considered the effect of major impacts on plate motion in his book [4, see chapters 6–8], but does not consider the lithospheric slip displacement introduced in this paper.

We consider an asteroid impact process which is known to be a low, but greater-than-zero probability event [10, 11]. We assume that the asteroid impacts the plate at an angle of incidence θ with respect to the surface of the plate. For a perpendicular angle of incidence $\theta = 90^\circ$, the impact will cause damage to the crust/lithosphere, with no slip displacement.

In addition, we consider an asteroid of mass m and speed v with respect to the plate which is assumed to initially be at rest. Then the asteroid's momentum in the plate's local plane is given by

$$p = mv \cos\theta \,. \tag{5}$$

When the asteroid collides with the plate, the collision's applied force impulse is given by

$$F_a = \frac{\Delta p}{\Delta t} \tag{6}$$

where $\Delta p = mv \cos\theta$ is the change in momentum of the plate assuming it is initially at rest and $\Delta t = \Delta t_p$ is the time interval for the momentum transfer, which is much shorter than Δt_c, the duration of the collision. Thus

$$F_a = \frac{mv \cos\theta}{\Delta t_p} \,. \tag{7}$$

Combining (2) and this equation, the slip condition for a plate slip displacement to occur in the direction of the collision as a result of the applied force overcoming the force of static friction becomes

$$\frac{mv \cos\theta}{\Delta t_p} > \mu_s Mg \,. \tag{8}$$

The variables on the L.H.S. are dependent on the characteristics of the asteroid and the collision, while those on the R.H.S. are dependent on the plate impacted.

We return to our sample calculation for the North American plate of section 2.1 to obtain an order-of-magnitude estimate of the effect under consideration. We consider a colliding asteroid of diameter $d \sim 20$ km, mass $m \sim 2 \times 10^{16}$ kg, $v \sim 30$ km/s, and use an angle of incidence $\theta = 45^\circ$ [12, 13]. Then substituting into (8) and using (3), we obtain slip condition

$$\frac{4 \times 10^{20}}{\Delta t_p} > 5 \times 10^{22} \, \text{N} \,, \tag{9}$$

which is dependent on the momentum transfer time. We consider three momentum transfer times: 1 s, 1 ms and 1 μs:

$$\text{for } \Delta t_p = 1\,\text{s}, \qquad 4 \times 10^{20}\,\text{N} \ngtr 5 \times 10^{22}\,\text{N}\,,$$

$$\text{for } \Delta t_p = 1\,\text{ms}, \qquad 4 \times 10^{23}\,\text{N} > 5 \times 10^{22}\,\text{N}\,, \qquad (10)$$

$$\text{for } \Delta t_p = 1\,\mu\text{s}, \qquad 4 \times 10^{26}\,\text{N} > 5 \times 10^{22}\,\text{N}\,.$$

Price [4, p. 171] notes that two stress waves are generated at the point of impact, one in the asteroid rocks and one in the plate rocks. These he estimates to each propagate at about 8 km/s, which points to a momentum transfer time in the ms range.

The slip condition is satisfied for the two shorter collision times (1 ms and 1 μs), but not for the longer one (1 s). Thus we find that lithospheric slip displacements are possible in plate tectonics under certain asteroid impact conditions. These are found to depend on the mass of the plate, the mass, velocity and angle of incidence of the asteroid, and the duration of the momentum transfer. The probability of a lithospheric slip displacement would be much higher for larger asteroids. We now investigate some of the details of the resulting motion of lithospheric slip displacements under asteroid impact conditions.

3 The conservation of energy model

In the previous section, we have considered the force model underlying lithospheric slip displacements in plate tectonics. In this section, we examine the motions resulting from the law of conservation of energy.

Before the collision, the energy of the plate-asteroid system, assuming the plate is at rest, is given by the kinetic energy of the incoming asteroid

$$E_i = \frac{1}{2}\,mv^2\,, \qquad (11)$$

where the variables are as defined previously. The collision is completely inelastic and the kinetic energy of the colliding body is transferred to the plate. In addition, energy is lost in the fracas, cratering and deformation of the plate as a result of the collision. After the collision, the energy of the plate-asteroid system is given by

$$E_f = \frac{1}{2}\,(M + m)\,V^2 + E_{rel}\,, \qquad (12)$$

where $m \ll M$, V is the velocity of the plate after the collision, and E_{rel} is the non-kinetic energy released in the collision. It should be noted that the slip of the plate as a result of the collision will reduce the non-kinetic energy E_{rel} released in the collision as the plate will yield to the asteroid and its motion will absorb a proportion of the collision energy.

To simplify our calculations, from the conservation of energy equation $E_i = E_f$, we write

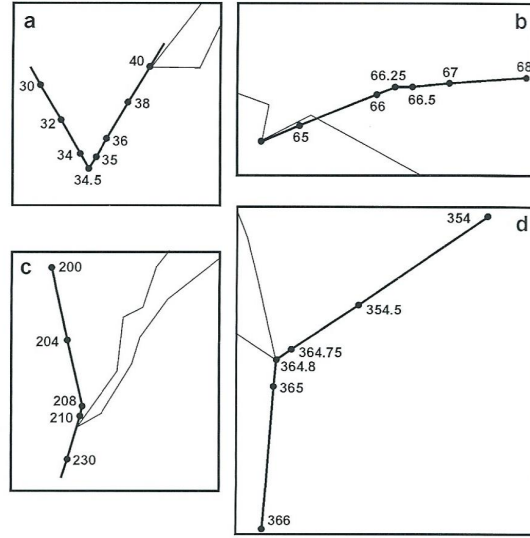$$\frac{1}{2}\,(M + m)\,V^2 = \frac{1}{2}\,\epsilon\,mv^2\,, \qquad (13)$$

Fig. 2: Figure 6.1 from Price [4, p. 196], caption: "Tracks related to four known impact structures. Examples a, b and c are of *certain* impacts. (a) is that for the Popigai crater, diameter 100 km and age 34.6 Ma. (b) is that of Chicxulub, diameter about 200 km and age 66.25 Ma. (c) is that of Manicouagan, diameter 100 km and age 208 Ma. (d) is for a 'near certain' CNCF [Central Nevada Circular Feature] impact, diameter about 220 km and date 364.8 Ma." Note the significant change in speed of the plate in example (c) after the change in direction.

where $\epsilon \le 1$ is the proportion of the initial energy transformed into kinetic energy of the plate, with the rest released as non-kinetic energy. Solving for V, we obtain

$$V = \sqrt{\epsilon\,\frac{m}{M}}\,v \qquad (14)$$

where we have neglected m in the term $(M + m)$.

We wish to calculate the distance that will be covered by the plate as a result of the lithospheric slip displacement. From (4) of the force model of section 2.1, we know that the plate will be subject to a constant deceleration $a = -\mu_k\,g$. We can thus use the dynamic equation

$$V_f^2 = V_i^2 + 2as \qquad (15)$$

where V_i is given by (14) and $V_f = 0$ when the plate stops moving. Solving for the distance s, we obtain

$$s = \frac{\epsilon}{2\,\mu_k g}\,\frac{m}{M}\,v^2\,. \qquad (16)$$

Using the values used in the sample calculation for the North American plate of section 2.1 and $\epsilon = 1$ implying that most of the energy is available as kinetic energy, we get an initial plate velocity $V_i = 32$ m/s from (14) and a lithospheric

slip displacement $s = 420$ m from (16). For $\epsilon = 0.5$ implying that 50% of the collision energy is available as kinetic energy, we get an initial plate velocity $V_i = 22$ m/s from (14) and a lithospheric slip displacement $s = 210$ m from (16). These values would be evident in the analysis of tectonic plate movements in the case of observed sudden changes in direction and/or speed of plate motions. In Fig. 2, we give examples from Price [4, Figure 6.1, p. 196] of plate tracks likely caused by lithospheric slip displacements resulting from asteroid impacts.

4   Discussion and conclusion

In this paper, we have considered simple models for order-of-magnitude proof-of-concept model calculations for lithospheric slip displacements in plate tectonics. We have obtained physically realistic results that provide an explanation for the observations:

- For a lithospheric slip displacement to occur, a force F_a must be applied to the lithospheric plate to overcome the force of static friction F_s that is holding it in place on top of the asthenosphere: $F_a > F_s = \mu_s M g$.

- The magnitude of the required applied force F_a can be generated in asteroid impacts. Lithospheric slip displacements are then possible under the following slip condition: $mv \cos \theta / \Delta t_p > \mu_s M g$. The asteroid impact condition is found to depend on the mass of the plate, the mass, velocity and angle of incidence of the asteroid, and the duration of the momentum transfer.

- The distance s that is covered by the plate as a result of the lithospheric slip displacement is given by $s = \epsilon m v^2 / 2 \mu_k g M$, under the action of a constant deceleration $a = -\mu_k g$, which explains observed sudden changes in direction and/or speed of plate motions as seen in Fig. 2.

The model calculations presented in this paper provide proof-of-concept evidence for lithospheric slip displacements in plate tectonics resulting from asteroid impacts. The model depends on many variables including the plates, asteroid and impact involved, and provides a framework to analyze such problems.

Many simplifications have been made that can lead to inaccuracies and complications, such as irregularities of the lithosphere and asthenosphere impacting the friction force, the proportion of collision energy being lost in the inelastic collisional process and not transformed into kinetic energy, *etc*. In addition, subsequent plate collisions resulting from the initial lithospheric slip displacement have to be analyzed for individual event conditions. Subsequent high-speed plate collisions could be a contributing factor to orogeny events resulting from violent plate collisions.

It should be noted that residual plate speeds, believed to be generated by mantle convection, are in the cm/annum range [4, p. 16]. Plates can thus be initially taken to be at rest in
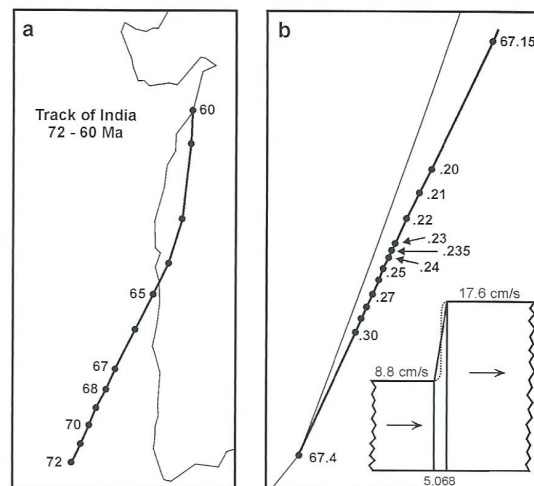

Fig. 3: Figure 6.7 from Price [4, p. 202], caption: "(a) Track of 'Bombay' over the period 72-60 Ma. It can be inferred from the distances between the points representing specific times that there was an abrupt change in velocity of the plate at about 67 Ma. (b) A detail of the track shown in (a) reveals that the velocity of plate motion doubled in a period which is assumed to be about 5000 years." There is an error in the units of the reported plate motion (cm/s instead of the correct cm/a) in the insert in (b).

the calculations in this paper. As Price [4, Figure 6.1, p. 196] notes, plate speed is changed along with direction in impact events. For example, he notes that the Manicouagan impact event (item (c) in Fig. 2) sped up the plate speed by a factor of 4 (in cm/annum), while for the others, the changes were -4-5% for item (a), 5-6% for item (b) and 11% for item (d).

The process of lithospheric slip displacement proposed in this paper would lead to a rapid change in plate direction and speed which would be followed by a change in residual plate speed in the cm/annum range, likely arising from the follow-on plate collisions that occur following a lithospheric slip displacement. The change in direction and the change in speed depend on the particulars of the impact event and cannot be easily calculated, requiring a detailed analysis of the particular impact event of interest.

Price, using the Atlas Version 3.3 software system [4, p. 192] to analyze plate track changes, has studied the Indian Deccan Traps geological structure that he attributes to a major impact event at 67.23 Ma which resulted in a change in plate direction and speed from 8.8 cm/a to 17.6 cm/a, to try to better understand the timeframe involved for the change in plate speed. In Fig. 3, we show the figure from Price [4, Figure 6.7, p. 202] in which he narrowed down the interval of plate speed change to less than 5 000 years (as shown in Fig. 3b). As he mentions, the rise-time would likely follow the S-curve shown in the insert in Fig. 3b, hence over a time interval shorter than 5 000 years. In his analysis, he attributes

a time for acceleration and for deceleration before the plate settles in its new residual plate speed (the short horizontal portions before and after the vertical portion of the S-curve shown in the insert in Fig. 3b).

The change in plate direction and speed is thus extremely short in geologic time. The model suggested in this paper shows that the time duration of the lithospheric slip displacement would indeed be very short both in geologic and in actual event time. This model provides an explanation for the abrupt changes of rate and direction of plate motion observed in the geological record. It provides a physical and mathematical framework for the analysis of lithospheric slip displacements in plate tectonics.

Received on July 7, 2020

References

1. Frisch W., Meschede M., Blakey R. C. Plate Tectonics: Continental Drift and Mountain Building. Springer, CH, 2011.

2. Livermore R. The Tectonic Plates Are Moving. Oxford University Press, Oxford, 2018.

3. Murphy J. B., van Andel T. H. Plate Tectonics. Encyclopaedia Britannica, www.britannica.com/science/plate-tectonics, accessed June 2020.

4. Price N. J. Major Impacts and Plate Tectonics. Routledge, New York, 2001.

5. Sleep N. H., Fujita K. Principles of Geophysics. Blackwell Science, Malden, MA, 1997.

6. Poirier J.-P. Introduction to the Physics of the Earth's Interior. Cambridge University Press, Cambridge, 1991.

7. Lithgow-Bertelloni C. and Richards M. A. The dynamics of Cenozoic and Mesozoic plate motions. *Reviews of Geophysics*, 1998, v. 36, 27–78.

8. Coefficient of Friction Table. roymech.org/Useful_Tables/Tribology/co_of_frict.html, accessed June 2020.

9. Static Friction Coefficients. structx.com/Material_Properties_007.html, accessed June 2020.

10. Berggren W. A. and van Couvering J. A., eds. Catastrophes and Earth History: The New Uniformitarianism. Princeton University Press, Princeton, NJ, 1984.

11. Verschuur G. L. Impact! The Threat of Comets and Asteroids. Oxford University Press, Oxford, 1996.

12. List of exceptional asteroids. en.wikipedia.org/wiki/List_of_exceptional_asteroids, accessed June 2020.

13. Near-Earth objects. en.wikipedia.org/wiki/Near-Earth_object, accessed June 2020.

Symmetry Breaking Model of Volume Pulsating Walking Droplets

Gianpaolo Bei[1], Davide Passaro[2]

[1]Darwin Secondary High School, Via Tuscolana 00181, Rome, Italy. E-mail: capo.gian@libero.it
[2]Department of Statistical Sciences, Rome La Sapienza, Piazzale Aldo Moro 5 00185, Rome, Italy.
E-mail: davide.passaro@uniroma1.it

In this article, we propose a generalized model of dynamic of extended pulsating walking droplets. In the first section, we provide a brief overview of the open problems of walking droplets. In the second section, we analyze some critical issues of the general stroboscopic models. In the third section, we elaborate our proposal of a generalized model of pulsating droplets. Finally, we suggest a link between walking droplets dynamic and the acoustic gravity wave induced on the surface of the vibrating bath.

1 Open problems of walking droplets

In the last fifteen years, the classical study of hydrodynamical Faraday waves has attracted great renewed interest since the discovery of walking droplets and the more general discovery of hydrodynamical pilot wave models [1,2]. Notwithstanding that many papers have cleared and rationalized a lot of phenomena with similarities to quantum mechanics (wave/particle duality, discrete orbits, tunnelling effect, statistical properties, *etc* [3, 4, 6, 7]), we propose that something is missing in the general approach to these issues. For example, as far as the authors are aware, there are no papers which explore an hydrodynamic analogue of the Planck law or of the de Broglie hypothesis or an analogue of the Born statistical interpretation of the wavefunction.

In particular, we propose that the role of volume pulsations and bath deformation may be caused by the impact of the droplet and its influence on the dynamic and conservation of the momentum of the global systems should be explored.

From our point of view, the problem of defining the total momentum of the particle-wave coupled system is deserving of closer study, since we believe that not only the momentum of droplet and the vibrating bath must be considered, but also those of the surface acoustic wave produced by the impact and of the vibrating borders of the vessel.

Our proposal is that the symmetry breaking force of the transition from the bouncing to the walking regime could be due to space asymmetries of one between the deformation of droplet, or the bath's deformation or the acoustic wave pattern or to an asymmetric vibration of the border.

In fact, the actual modellization of the transition between the bouncing regime and the walking regime is based on the surface orography of the vibrating bath, but this model does not yet justify the mechanism by which the surface has a broken symmetry and moreover it assumes that the droplet is punctiform. We assume that the surface bath geometry asymmetry is caused by an acoustic gravity wave and not just by a surface gravity wave [20].

Furthermore, at present there is no model that has a frequency dependent broken symmetry mechanism.

Finally, although there are some experimental studies of the droplet volume pulsation, presently we lack a model that tries to implement this experimental fact. In the following section, we analyze some critical aspects of the stroboscopic model which we believe are yet to be explored.

2 Critical aspects of stroboscopic models

The stroboscopic model of Bush-Molaceck and the generalized integral model of Oza [9, 11, 13] has been till now the most successful and most used model to rationalize walking droplets.

The two major hypotheses on which it is based are the following [5]:

1) The bath height oscillations are described by standing monochromatic waves.

2) The bath Faraday wave field is resonant with the bouncing oscillations (the mode is (2,1)).

The efforts to improve and generalize this model are stimulated by the desire to extend it to multiple droplets dynamics and to describe more accurately the spatio-temporal decay of the bouncing induced Faraday waves.

In the following, we will describe some other hypotheses which we consider need to be better justified and maybe generalized.

The general approach to describe droplet-bath dynamics is to separate the horizontal and the vertical dynamics during flight; on the contrary, we believe that if we want to describe more accurately the real spatio-temporal extended impact between the drop and the bath, we have to consider the successive volume oscillations of the droplet and the acoustic waves beneath the surface bath.

In fact, they persist after the impacts and therefore implement a dynamical memory dependent coupling which moreover hides some energy and momentum whose conservation may be deepened.

The first stroboscopic model [9] contained discrete sums of Bessel functions describing the wavefield and used in the trajectory equation averaged over the bouncing period:

$$m\ddot{x}_i + D\dot{x}_i = -mg\, S\left(h_i(x_i, t)\right) \nabla h_i(x_i, t) \qquad (1)$$

where D is the drag coefficient, h is the bath height and $S = \sin \Phi$ is the impact phase which is dependent on the mean phase of the wave during the drop contact.

In particular, this model assumes that the height of the vibrating bath is given by a linear superposition of n circular waves each one generated by the drop impact described by the following relation:

$$h(\mathbf{x}, t_N) = \sum_{-\infty}^{N-1} \frac{A e^{-(x-x_n)/\delta}}{|x - x_n|^{-1/2}} \, e^{-(t_N - t_n)/\tau} \cos(k_F |\boldsymbol{x} - \boldsymbol{x}_n| + \Phi) \,. \quad (2)$$

Recently, some authors [12] have proposed generalizations based on the mean wave field, but all the generalizations are based on the hypothesis of instantaneous and punctiform gradient of the surface wave slope and are aimed to rationalize the wavelike statistics of irregular unstable orbits.

Finally, we wish to note that thus far we lack a self-consistent explanation of the origin of the symmetry breaking force and its associated horizontal momentum transfer.

In the following section, we want to discuss a proposal which attempts to overcome these difficulties and to connect this problem to the search of an energy minimization principle which could explain the main features of the walking droplets stable orbits.

3 Generalized stroboscopic model

We propose to generalize the stroboscopic model by introducing a horizontal force which depends on the frequency and the volume pulsation of the droplet. In particular we implement a memory dependent force taking into account the previous volume oscillation.

Given an horizontal plane of the non-vibrating bath represented by x and y, our generalized symmetry breaking force starting from [10] is the following:

$$\vec{F}_{xy} = \int_{t-t_0}^{t} \nabla p \, \dot{V} \, d\tau = m \frac{\Delta \vec{v}_{xy}}{\tau_0} \quad (3)$$

where:

- t_0 is the impact time of the droplet with the bath and it is the inverse of the frequency of the volume pulsation;
- V is the volume of the droplet and $\dot{V}$ is the derivative with respect to time;
- ∇p is the gradient of the bath pressure wave.

This force disappears when $t_0 = 0$, while it converges to that one of the stroboscopic model when the frequency of the pulsation is 0.

Our proposal assumes that this force is present only during the impact and that the pressure on the droplet is due to the potential gravitational energy of the deformed bath.

In fact, differently from the Bush-Molacek model, the real geometrical profile of the vibrating path during the impact is no more sinusoidal. The bath absorbs elastic energy from the bouncing droplet during the impact and consequently it is deformed.

The height difference between the sinusoidal profile and the modified profile gives the potential energy to the deformed droplet.

Our hypothesis is that the pressure p and the height difference are given by the following formula derived from the theorem of conservation of the fluid energy:

$$p + \rho_{bath} \, g_{eff} \, \Delta h = \cos t \quad (4)$$

where ρ is the bath density, g_{eff} is the same used in the stroboscopic model [13] (also denoted as g_*) and p is the pressure induced in the bath after the droplet's impact.

This equation can be generalized since the external vibrating force continuously adds energy to the bath:

$$p + \rho_{bath} \, g_{eff} \, \Delta h = \alpha(t) \quad (5)$$

where $\alpha(t)$ is a periodic function dependent on the oscillatory force and on the volume deformation; Δh is the variation of the harmonic oscillation of the height of the bath caused by the impact of the droplet.

The introduction of this force (which is present only during the impact) requires a generalization of the horizontal dynamics of the walking droplet. Moreover we continue to assume the usual vertical periodic dynamic of stroboscopic model.

If during the impact, we apply to Newton equation (3) using the formalism of the finite difference instead of the derivative, the gradient operator to (5), we arrive at the following model (since we assumed that α depended only on the time t):

$$m \frac{\Delta \vec{v}_{xy}}{\tau_0 + T_F} + \int_{t-\tau_0}^{t+T_F} D \frac{\vec{v}_{xy} \, \dot{V}}{\Delta V} =$$
$$= - \int_{t-\tau_0}^{t+T_F} \rho \, \nabla(g_{eff} \, \Delta h) \, \dot{V} \, d\tau \quad (6)$$

where:

- T_F is the inverse of the Faraday frequency of the vibrating bath;
- the instantaneous acceleration used by the stroboscopic model has been substituted by the finite difference variation of the velocity during the impact time τ_0;
- $V(t)$ is the time dependent volume pulsation of the droplet that can be assumed to be described by the following formula:

$$V(t) - V_0 = V_0 \cos(\omega t) \, e^{-\lambda t}$$

an exponential decay of an harmonic oscillation with ω the frequency of droplet self-mode oscillation and λ the time decay coefficient;

- g_{eff} is the asymmetric effective gravity dependent on the local frequency given by the following relation:

$$g_{eff} = \gamma \, \sin[2\pi f(x) \, t]$$

where $f(x)$ is the space-dependent local frequency caused by the asymmetric acoustic wave interference process not considering the dissipation;

- the first integral is a temporal average of the drag force over the past volume pulsation of the droplet of the drag force;

- the second integral has been obtained from (4);

- the gradient in the last integral is due to space asymmetry of the effective gravity of the bath, which we hypothesized could be associated to a space dependence of the bath vibrating frequency.

It is interesting to show that it is possible to recover the main aspect of the stroboscopic model in the following way:

- the first term $m \frac{\Delta \vec{v}_{xy}}{\tau_0 + T_F}$ gives the usual discretized acceleration when the impact time τ_0 goes to zero;

- the second term $\int_{t-\tau_0}^{t+T_F} D \, \frac{\vec{v}_{xy} \dot{V} d\tau}{\Delta V}$ becomes the dissipative term during the flight when the impact time goes to zero;

- the second member is able to reproduce the slope gradient term $\overline{-F(t) \, \nabla h(x_p, t)}$ introduced by Bush et al [11] when τ_0 tends to zero and applying the gradient to (4); in the stroboscopic model the effective gravity is assumed to be space-independent differently from our model.

This model, of course, contains a hidden variable that is the space-dependent frequency vibrating of the bath. This variable allows to fit the numerical model in order to be in agreement with the stroboscopic model, but could be deduced by coupling (6) with another law that relates the pressure with the volume pulsation, assuming that $\alpha(t)$ of (5) is proportional to the second time derivative of the droplet volume [16].

On the contrary to the stroboscopic models, we don't make any ad hoc assumption on the geometric pattern of the surface wave since we think that it should be deduced by investigating experimentally the acoustic spectrum of the surface acoustic gravity wave.

Among many ad hoc and arbitrary hypotheses, we think that a simple option could be the sound emission law taken from [16]:

$$\phi = -\frac{\dot{V}(t)}{4\pi r} \tag{7}$$

where r is the position of a point with respect to the initial impact of the droplet and ϕ is the usual velocity potential of the bath that is related to the effective gravity described by the following formula:

$$\nabla \dot{\phi} = a \cdot g_{eff} \tag{8}$$

with a a dimensional constant.

Finally, we assume that the oscillating acoustic pressure perturbation and the acoustic velocity field obey the following equations of motion:

$$\rho \, d_t \vec{v} = -\nabla p, \quad \beta \, d_t p = -\nabla \cdot \vec{v}. \tag{9}$$

where ρ is the density of the bath, p is the acoustic pressure, $d_t \vec{v}$ is the convective temporal derivative of the moving fluid, β is the inverse of B the bulk modulus of the acoustic pressure wave [15]; this is a self-consistent system of partial differential equations which determines the coupled dynamic of the system.

This choice is motivated by the link between an oscillating volume and the generation of an acoustic spin wave in a fluid as described in [19]. We suggest that it could be interesting for the experimental researcher to study the change of the acoustic spectrum during the transition from the bouncing regime to the walking regime and could be an operative way to verify or, eventually, falsify the general model proposed.

4   Conclusions

We have studied the problem of the origin of the symmetry breaking force that causes the asymmetry of the wave pattern of vibrating bath. We propose a generalized stroboscopic model of an extended and deformable walking droplet.

In particular, our proposal is based on the hypothesis that each bounce generates an acoustic gravity on the surface and its asymmetric reflection causes a space dependent bath vibrating frequency.

Recently a new class of walking droplets, called superwalkers, have been discovered [21]. These new observations show a strong correlation between the volume of the droplet and the duration of the impact with the velocity of the walking droplet. This property may be interpreted as an indirect confirmation of our hypothesized coupling between the volume deformation and the droplet dynamic.

We hope that our approach will stimulate more extensive experimental research on the energy of the global system (droplet and vibrating bath).

In particular we think that all the models lack an explanation of the role of the energy and its non-conservation and minimization on the discrete orbit of the walking droplets; in fact, the dynamics of stroboscopic models of walking droplets is based on empirical models and not on a general variational principle of this peculiar dissipative system.

Our insight is that the energy and the impulse of the horizontal motion of the walking droplets are associated to the volume oscillation and the deformation of the bath which induces an acoustic gravity wave with momentum and energy.

Furthermore, our opinion is that the volume oscillation would induce density waves in the bath whose turbulence could be explained by onset of turbulence as studied by Francois et al [17], whose origin could be caused by helicoidal under surface sound waves.

We think that this hidden energy due to volume pulsation could be experimentally investigated studying the relation with the momentum of the under bath acoustic wave; it is fascinating to speculate that the law behind this could be given by an acoustic hydrodynamic de Broglie-like relation inherent to the energy of droplet volume pulsation:

$$m \, \Delta v_{xy} = H \, k \qquad (10)$$

where the first member refers to the kinematic momentum of the droplet, and the second member is related to the acoustic wave momentum with H the hydrodynamic analogue of the Planck constant.

Finally, we think that it could be useful to explore experimentally the possibility to induce the transition from the bouncing regime to the walking regime, making oscillating the vessel keeping constant the frequency and the modulus of the shaker vertical acceleration; we expect that there will be a critical phase transition in a preferred direction from the bouncing to the walking regime.

Received on June 26, 2020

References

1. Couder Y., Protière S., Fort E., Boudaoud A. Dynamical phenomena - Walking and orbiting droplets. *Nature*, 2005, v. 437 (7056), 208.

2. Fort E., Eddi A., Boudaoud A., Moukhtar J., Couder, Y. Path-memory induced quantization of classical orbits. *PNAS*, 2010, v. 107 (41), 17515–17520.

3. Labousse M., Oza A. U., Perrard S., Bush J. Pilot-wave dynamics in a harmonic potential: Quantization and stability of circular orbits. *Physical Review E*, 2016, v. 93, 033122.

4. Couder Y., Fort E. Single-Particle Diffraction and Interference at a Macroscopic Scale. *Physical Review Letters*, 2006, v. 97 (15), 154101.

5. Turton S. E., Couchman M. M. P., Bush J. W. M. A review of the theoretical modeling of walking droplets: Toward a generalized pilot-wave framework. *Chaos*, 2018, v. 28, 096111.

6. Eddi A., Moukhtar J., Perrard S., Fort E., Couder Y. Level Splitting at Macroscopic Scale. *Physical Review Letters*, 2012, v. 108, 264503.

7. Nachbin A., Milewski P. A., Bush J. W. M. Tunneling with a hydrodynamic pilot-wave model. *Physical Review Fluids*, 2017, v. 2 (3), 034801.

8. Oza A., Harris D., Rosales R., Bush J. W. M. Pilot-wave dynamics in a rotating frame: On the emergence of orbital quantization. *Journal of Fluid Mechanics*, 2014, v. 744, 404–429.

9. Moláček J., Bush J. W. M. Drops walking on a vibrating bath: Towards a hydrodynamic pilot-wave theory. *Journal of Fluid Mechanics*, 2013, v. 727, 612–647.

10. Leighton T. G., Walton A. J., Pickworth M. J. W. Primary Bjerknes forces. *European Journal of Physics*, 1990, v. 11 (1).

11. Oza A. U., Rosales R. R., Bush J. W. M. A trajectory equation for walking droplets: Hydrodynamic pilot-wave theory. *Journal of Fluid Mechanics*, 2013, v. 737, 552–570.

12. Durey M., Milewski P. A., Bush J. W. M. Dynamics, emergent statistics, and the mean-pilot-wave potential of walking droplets. *Chaos*, 2018, v. 28 (9), 096108.

13. Turton S. E., Couchman M. M. P., Bush J. W. M. A review of the theoretical modeling of walking droplets: Toward a generalized pilot-wave framework. *Chaos*, 2018, v. 28, 096111.

14. Tambasco L. D., Bush J. W. M. Exploring orbital dynamics and trapping with a generalized pilot-wave framework. *Chaos*, 2018, v. 28 (9), 096115.

15. Bliokh K. Y., Nori F. Spin and orbital angular momenta of acoustic beams. *Phys. Rev. B*, 2019, v. 99, 174310.

16. Landau L. D., Lifshitz E. M. Fluid Mechanics. Butterworth-Heinemann, Oxford, 1987.

17. Francois N., Xia H., Punzmann H., Ramsden S., Shats M. Three-Dimensional Fluid Motion in Faraday Waves: Creation of Vorticity and Generation of Two-Dimensional Turbulence. *Phys. Rev.*, 2014, 021021.

18. Filatov S. V., Parfenyev V. M., Vergeles S. S., Brazhnikov M. Y., Levchenko A. A., Lebedev V. V. Nonlinear Generation of Vorticity by Surface Waves. *Phys. Rev. Lett*, 2016, v. 116, 054501.

19. Burns L., Bliokh K. Y., Nori F., Dressel J. Acoustic versus electromagnetic field theory: scalar, vector, spinor representations and the emergence of acoustic spin. arXiv: class-ph/1912.10522.

20. Kadri U., Stiassnie M. Generation of an acoustic-gravity wave by two gravity waves, and their subsequent mutual interaction. *Journal of Fluid Mechanics*, 2013, v. 735, R6.

21. Valani R. N., Slim A. C., Simula T. Superwalking Droplets. *Phys Rev Lett*, 2019, v. 123 (2), 024503.

The Ambiguity of Celestial Dynamics

Alfred N. van Hoek

Formerly: Department of Neurology, University of Utah, 36 South Wasatch Drive, SMBB 4322, Salt Lake City, UT 84112-5001.
E-mail: alfred.vanhoek@hsc.utah.edu; vanhoek@mac.com. ORCID 0000-0003-2724-6855.

Since the discovery of stellar aberration, human perception failed to recognize the fundamental property of motion parallax to recover the depth of the universe. Stellar aberration, the motion of the fixed stars in the perceived direction of Earth's motion, is the essence of reversed perspective [Purves D., Andrews T. J. *Proc. Natl. Acad. Sci. USA*, 1997, v. 94, 6517–6522]. The true-to-reality perception requires a finite-radius celestial sphere, which functions as a non-inertial frame of reference; its coordinates along the line-of-sight describe a Coriolis circulation at a parallax distance of 58.13 light-days.

1 Introduction

The perception of a three-dimensional universe projected onto a two-dimensional projection surface of the celestial sphere is or becomes equivocal, because the uncertainty [1] requires the distinction between illusion and veridicality. The recovery of the missing third dimension, the depth of field is prone to these two possibilities and may be best described by the Necker illusion [1], which switches between proximal and distal faces of a two dimensional representation of a cube. The spatial relationship of the proximal and distal faces can be ascertained by motion parallax so that the object moves laterally in relation to the background, thereby providing perspective or true-to-reality perception [1]. The illusionary perception, the distal and proximal faces of the cube are perceived to be front, respectively, back, associates with reversed perspective and motion parallax fails; the background appears to rotate in the direction of motion. When considering celestial sphere grids at a finite and at an infinite distance, the motion of Earth around the Sun will cause parallax of the proximal grid in annual fashion with respect to the fixed stars. However, perceiving the sphere surface as the distal grid, the motion of the Earth will cause negative parallax of the stars, which is known as stellar aberration [2]. Special relativity proposed that space contraction in the direction of motion is a logical consequence of the universal constant, the finite speed of light. Among other laws of motion, it proclaimed the law of stellar aberration [3], which as it stands is incompatible with the finite-radius celestial sphere. To discern illusion from reality, we address the intricacy of celestial sphere radii (finite or infinite) in this thesis.

2 Celestial sphere considerations

The nature of a celestial sphere centered on Earth with a fixed orientation (Fig. 1) and the apparent alignment with the fixed stars suggests a stationary frame of reference. However, the discovery of stellar aberration in the direction of Earth's motion was a surprising phenomenon because a fixed point at the firmament should not cause any measurable displacement. Mathematically, centering the celestial sphere to Earth or to an arbitrary planet in a fixed configuration, as shown in Fig. 1,

two rotational frame of references need to be reconciled with. The anti-clockwise planetary orbit drives the celestial sphere into rigid body circulation, subjecting the coordinates, emanating from the centre towards the surface of the sphere along the spindles, to the nonzero curl of the velocity field **u**,

$$\boldsymbol{\xi} = \nabla \times \mathbf{u}, \tag{1}$$

also known as the vorticity $\boldsymbol{\xi}$. Because the rotation occurs in the x-y plane, the vertical component ζ is nontrivial,

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \tag{2}$$

where u and v are the velocity components of the planetary orbital velocity $\Omega \times R$, i.e.

$$\begin{aligned} u &= -\Omega y \,\hat{\mathbf{x}} \\ v &= \Omega x \,\hat{\mathbf{y}}. \end{aligned} \tag{3}$$

Substitution of (3) into (2) leads to the identity

$$\zeta = 2\Omega. \tag{4}$$

This means that the unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$ (Fig. 1, left panel) are locked to the orbital period of the planet, i.e, the vector $\hat{\mathbf{x}}$ is facing the rotational axis Ω. To steady the sphere in a fixed orientation requires a clockwise turning about its centre, which orients it in the stationary position (Fig. 1, right panel). This clockwise turning does not nullify the vorticity field of planetary motion, defining the celestial sphere system as a non-inertial frame of reference. Fig. 2A is a graphical representation of a celestial sphere centered on the planet with radial distance equal to the orbital radius of the planet. An arbitrary spindle, from the centre of the planet to the surface of the sphere, marked as 1-1, 2-2, 3-3, 4-4, 5-5 (Fig. 2A), represents the fixed line-of-sight towards the firmament and will describe an anti-clockwise circular trajectory. Divergent light-rays (exemplified in Figs. 2A, 2B) from a star to a receiver become convergent lines from receiver to the source as if they are parallel lines that vanish in perspective (Fig. 2A,
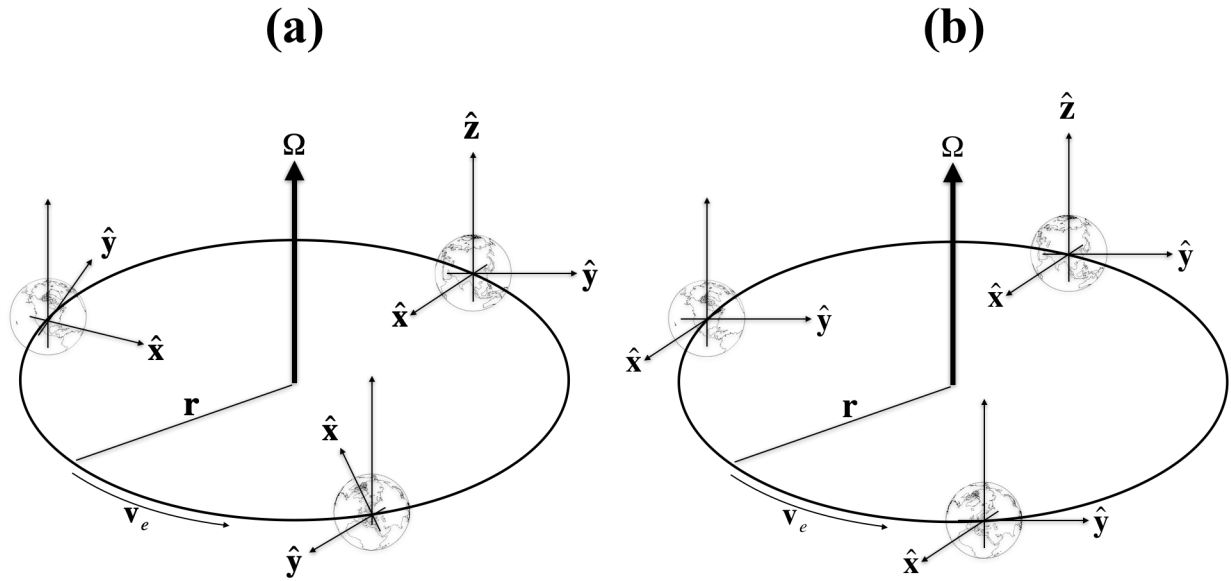
(a)

(b)

Fig. 1: **Coordinate systems.** An orbiting coordinate system around the stationary Sun is period locked (a).The unit vectors maintain a fixed orientation (b) by the clockwise annual spin of the celestial sphere about its axis $\hat{\mathbf{z}}$ to oppose the counterclockwise spin-orbit locking about Ω. The origin of the celestial sphere moves at constant angular velocity about the axis Ω with a fixed distance $\mathbf{r}$ from the centre of an inertial frame of reference.

2B), which also should hold when the light rays are truly parallel. This contrasts the divergent spindles of the celestial sphere that never can become parallel and cannot vanish in perspective as true parallel lines do[*]. Thus, with respect to the spindle of the celestial sphere, the direction to a star changes, forming different angles with the chosen spindle. The orbital trajectory of a spindle occurs further out in space when the radius of the celestial sphere is increased. The completion of a full planetary orbit of the sphere results in an imaginary Lissajous figure that is produced by a hula-hoop mechanism of the celestial sphere. The width of the donut-shaped Lissajous equals the diameter of the planetary orbit. This imaginary orbit, composed by the time-dependent endpoint of a single spindle, is thus formed by a bundle of parallel lines (Fig. 2C). If the radial distance of the spindle is increased we should expect the imaginary orbit to vanish in perspective. But the angling of the telescope as observed [2] appears to be a requirement to adjust the celestial sphere coordinate system to steady the stars [this thesis], suggesting a finite-radius celestial sphere. We then could conclude, from whatever direction observations are made, that the velocity field of a planet creates an imaginary orbit of a coordinate spindle about a star representing the Sun. The imaginary Sun is then, alike the endpoint of the celestial sphere spindle, located at the celestial sphere surface, exemplifying the imaginary Sun-Earth orbit system at a finite distance. This "kinematic optical" ef-

fect at a distance, the frame-dependent Coriolis circulation, is what aberration of light may represent and could be an equivalent to Snell's law.

3 The finite-radius celestial sphere

Figs. 3A and 3B highlight the angling necessary to maintain the line-of-sight towards the perceived stationary imaginary Sun Q in the ecliptic and pole directions, respectively. The line-of-sight coincides with a spindle of the celestial sphere and as shown the color coded circle and matching color-coded spindle/radius defines the line EQ with length C, and is equal to the centre-to-centre distance of the imaginary orbit and its planetary orbit. Since the line-of-sight can be chosen at will towards a star or an invisible point of interest, the angling towards Q in the figures, is caused by the changing position of the celestial sphere anchored to the orbital motion around the Sun. The fixed distance C, i.e. EQ, suggests (Fig. 4) the Scotch yoke reciprocating motion where the orbital position of the planet, point E, changes the position of the centre of the celestial sphere with respect to point Q (cf. Fig. 3). From the viewpoint of O, Q will slide along the vertical axis that coincides with the line OQ. According to the cosine rule, we have

$$(EQ)^2 = (OE)^2 + (OQ)^2 - 2 \cdot OE \cdot OQ \cdot \cos\theta,$$

$$(OE)^2 = (EQ)^2 + (OQ)^2 - 2 \cdot EQ \cdot OQ \cdot \cos\phi,$$

where EQ is the equivalent of the crank rod length equaling C and OE represents the orbital radius R. Substitution of the

---

[*]Infiniteness of the celestial sphere is usually interpreted as if the spindles are parallel lines. It then may be practical given the centre of the celestial sphere would be everywhere [5].
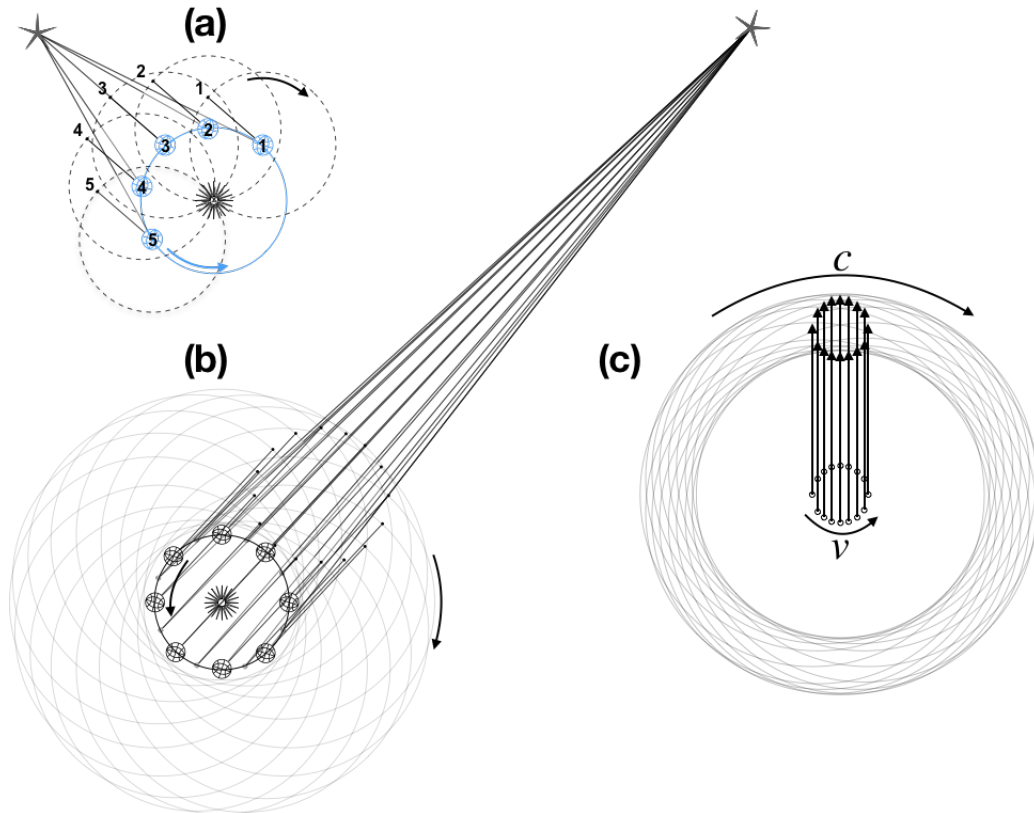
Fig. 2: **Celestial sphere radius and the velocity field.** Views of a planetary body rotating in a stationary frame of reference with the Sun as its centre. The celestial sphere centered on the planet with the same radius as the orbit (a) has a fixed orientation. The spindle of the celestial sphere at positions 1 trough 5 does not change in direction, while the direction to a star is dependent on the orbital motion of the planet. The increase of the celestial sphere radius (b) and (c) reduces the angling to a star at a finite distance. The line-of-sight to the surface of the celestial sphere (c) describes a circulation with the radius equal to the orbital radius. In perspective the subtended angle is equal to the parallax angle. To keep the celestial sphere in a fixed orientation, the axial clockwise rotation opposes the counterclockwise orbital motion of the celestial sphere. At the parallax distance the orbital velocity of the sphere surface is equal to the speed of light.

first cosine formula into the second cosine formula, replacing $(EQ)^2$, yields

$$OQ = R\cos\theta + C\cos\phi, \qquad (5)$$

where OQ represents the projection of the lines R and C with respect to the stationary reference frame of the Sun. The first term at the right hand side is the offset of point Q with respect to the Sun, caused by orbital motion. The second term at the right hand side describes the radial component of stellar aberration. The law of sines,

$$R\sin\theta = C\sin\phi, \qquad (6)$$

corresponds to the tangential component of stellar aberration. Substitution of (6) into (5) provides the combined form, independent of the subtending angle ϕ term, where the angle θ equals the angular velocity $\dot{\theta}$ of the planet at time t, yielding

$$OQ = R\cos\theta + \sqrt{C^2 - R^2\sin^2\theta}. \qquad (7)$$

Motion of point Q away or towards the Sun is the quintessence of the hula-hoop motion of the celestial sphere (Fig. 2 and 3), contributing to the decreasing and increasing parallax angle ϕ, when observing aberration in the direction of the plane of the ecliptic. The second term at the right hand side of (7), normalized to C,

$$\cos\phi = \sqrt{1 - \frac{R^2}{C^2}\sin^2\theta}, \qquad (8)$$

is complementary to (6). In terms of v and c, multiplying the radii R and C with the planetary angular velocity $v = \Omega \times R$ and $c = \Omega \times C$, the identities $vc^{-1} = \beta = RC^{-1}$ modify the above sine and cosine of ϕ (Eqs. 6, 8) to

$$\sin\phi = \beta\sin\theta \qquad (9)$$

and

$$\cos\phi = \sqrt{1 - \beta^2\sin^2\theta}. \qquad (10)$$
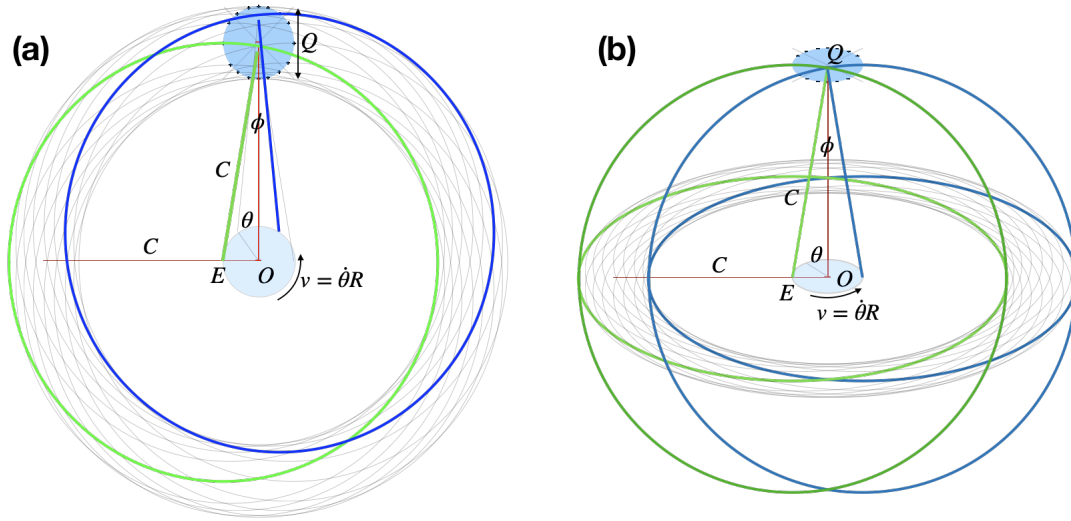
Fig. 3: **Geometry of negative parallax.** Point O represents the Sun and E is the position of the Earth in its orbit. Earth is the centre of the celestial sphere, the light blue circle is the Earth orbital plane and the dark blue filled circle is the traced-out orbital path of the Earth at distance C. (a): geometry of the celestial sphere in the ecliptic plane. (b): geometry of the celestial sphere towards the poles of the sphere.

Eq. (9) is reminiscent of the formula used to describe Snell's law, when referring to light or other waves passing through a boundary, which would be the celestial sphere surface that has a tangential velocity equal to the luminal speed c (Fig. 2C). The spindle EQ (Fig. 4) is the line-of-sight to and coincides with a light ray from the faraway fixed star of interest. The flight time of light from Q to E (see Sections 5 and 6) equals one radian of the orbit and is another property of the "kinematic optical" effect at a distance. The value of the aberration of light is defined when $\theta = \pm 90°$. The Pythagorean (9–10) becomes a right triangle. This condition is also equivalent to the line-of-sight when EQ is perpendicular to the ecliptic. The right triangle in terms of β is a Lorentz triangle with sides $1, \beta$ and $\gamma\beta$, where

$$\sin \phi = \beta,$$
$$\cos \phi = \sqrt{1 - \beta^2} = \gamma^{-1}. \tag{11}$$

The cosine term is identical to the reciprocal of the Lorentz gamma factor, i.e. $\gamma = \sec \phi$, which in terms of the special theory scales the Lorentz transformation matrix. Thus, the outcome of this treatise on the fixed and finite celestial sphere radius, leads to the same aberration of stellar light β but with opposite sign. In terms of the three-dimensional universe the stars are no longer perceived illusory and will behave veridically [1].

4 Transformation matrices

The general form of the special case (11), embodied by (9) and (10), also provides novel insight in which motion involves not only a change of β when considering the direction cosine, the line-of-sight, but also a change of the gamma-like factor (10). Given the radial vector $\boldsymbol{r}$ and time t, utilizing (9–10) instead of γ and β (11) as defined and used in the Lorentz transformation matrix [3], premultiplication of the vector $[t, \boldsymbol{r}]$ with the generalized and modified Lorentz transformation matrix containing the vorticity entries, i.e. (9) and (10),

$$\begin{bmatrix} t' \\ \boldsymbol{r}' \end{bmatrix} = \begin{bmatrix} \sec \phi & -c^{-1} \tan \phi \\ -c \tan \phi & \sec \phi \end{bmatrix} \begin{bmatrix} t \\ \boldsymbol{r} \end{bmatrix}, \tag{12}$$

results in

$$t' = \sec \phi \left(t - \boldsymbol{r} c^{-1} \sin \phi \right) \tag{13a}$$
$$\boldsymbol{r}' = \sec \phi \left(\boldsymbol{r} - ct \sin \phi \right). \tag{13b}$$

The derivation of the Lorentz transformation, matrix $\mathbf{L}$ involved the Galilean matrix, $\mathbf{G}$, and an assisting* or temporal matrix, $\mathbf{T}$. In generalized vorticity forms (cf. (12)) they become

$$\mathbf{G}_\odot = \begin{bmatrix} 1 & 0 \\ -c \sin \phi & 1 \end{bmatrix} \tag{14}$$

and

$$\mathbf{T}_\odot = \begin{bmatrix} \cos \phi & -c^{-1} \tan \phi \\ 0 & \sec \phi \end{bmatrix}. \tag{15}$$

Premultiplication of $\mathbf{G}_\odot$ with $\mathbf{T}_\odot$ gives the Lorentz matrix (cf. (12))

$$\mathbf{L}_\odot = \mathbf{T}_\odot \mathbf{G}_\odot. \tag{16}$$

If $\theta = \pm 90°$ (9–10), these matrices reduce to those Einstein derived. The subscript $\odot$ refers to a circular path with the line-of-sight along a spindle of the celestial sphere. Matrix $\mathbf{L}_\odot$ exemplifies the finite speed of light embodied by matrix

---

*The Lorentz matrix was heralded by the Zeitgeist of thence. The assisting system did not gain significance given its auxiliary status.
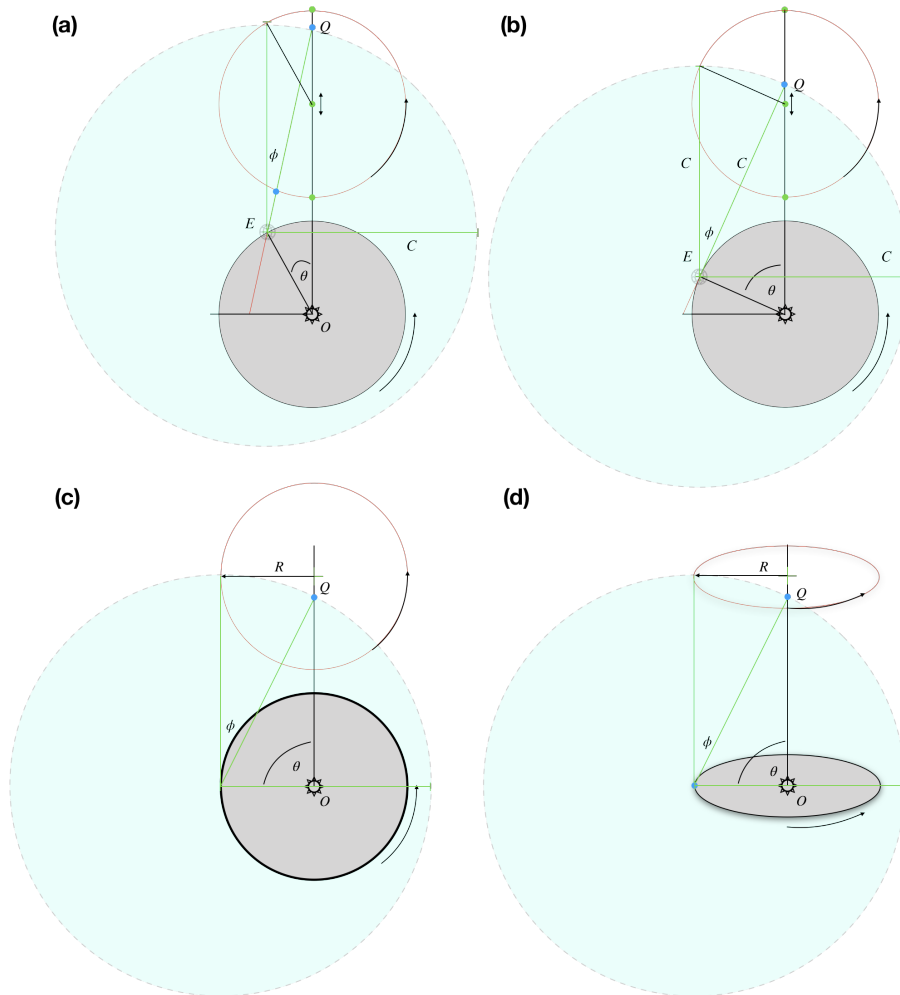
Fig. 4: **Scotch yoke reciprocating motion.** Earth's celestial sphere (light blue) with respect to the origin O, (the Sun) at (a), (b) and (c), affects the position of point Q, which is the imaginary Sun (see Section 2) as seen from Earth, E, sliding it up and down along the vertical line from O to Q. The line EQ is the line-of-sight (a celestial sphere spindle) with a fixed length. A complete revolution of Earth replicates the orbit of Earth (red colored). Earth's celestial sphere at (c) in the equatorial plane is replicated as (d) when the line-of-sight is towards the celestial pole.

$\mathbf{T}_\odot$, exemplifies the invariance of c and $\mathbf{G}_\odot$ exemplifies the Galilean transform. The location of a point in space and time is described by matrix $\mathbf{G}_\odot$; to detect this point by light requires a method, a transformation by using matrix $\mathbf{T}_\odot$. The identities

$$\mathbf{G}_\odot = \mathbf{T}_\odot^{-1}\mathbf{L}_\odot, \tag{17a}$$

$$\mathbf{T}_\odot = \mathbf{L}_\odot\mathbf{G}_\odot^{-1}, \tag{17b}$$

constitute a mechanism to transform a light-clock signal originating from a point and defined by matrix $\mathbf{L}_\odot$ to what will be an ordinary light-independent point in space and time defined by matrix $\mathbf{G}_\odot$. The true form of identifying an object is not perceived by light, which confirms Bradley's assessment 300 years ago or in other words these matrices correct for the delay of arrival time of light. The identities (17) suggest to convert light-signal based data to real data allowing ordinary addition of velocities and if necessary use the identity (16) to determine the Doppler effect.

5 The one radian of an orbit

The parallactic displacement of the coordinate system defined by β equaling $vc^{-1} = RC^{-1}$ for each of the planets of the solar system were calculated from their orbital radius R (Table 1, row 1) and period T (Table 1, row 2), yielding β, the radius of the celestial sphere C and the aberration angle $\phi = \arcsin\beta$ (Table 1, row 3). The celestial sphere radii increase with decreasing aberration angle, while the ratio of the celestial sphere radius C and period T of the planetary orbit

$$\frac{C}{T} = \frac{\Omega \times C}{2\pi} = \frac{c}{2\pi} \tag{18}$$

Table 1: The planets of the solar system are listed with their orbital radius R (row 1, [au]), period T (row 2, [year]) and the aberration angle ϕ [arcsec]. The duration of a light signal from the celestial sphere surface to the planet μ^{-1} (row 4, [day]) equals the one radian of the orbit, see (18). Based on the planet-Sun barycentre distance, bc (row 5, [km]), the solar orbit velocity v_{Sun} (row 6, [m/s]) was calculated using the planetary period T. The ratio of solar orbit velocity and the speed of light is provided in terms of an aberration angle ϕ_{bc} (row 7, [arcsec]), representing the planetary-specific celestial sphere of the Sun (see Section 7).

	Mercury	Venus	Earth	Mars	Jupiter	Saturn	Uranus	Neptune
R	0.39	0.72	1	1.52	5.2	9.54	19.2	30.1
T	0.241	0.615	1	1.88	11.9	29.5	84.1	164.8
ϕ	33.19	24.09	20.49	16.6	8.99	6.64	4.68	3.74
μ^{-1}	14.01	35.75	58.13	109.28	691.75	1,714.83	4,888.73	9,578.82
bc	10	265	445	74	742,465	408,110	12,585	230,609
v_{Sun}	0.007936	0.08577	0.08955	0.007786	12.467	2.759	0.2967	0.2786
ϕ_{bc}	0.000005	0.00006	0.00006	0.000005	0.0086	0.0019	0.0002	0.0002

is a constant, $10\,066.61$ au/year, i.e. one radian of the orbit, which equals the duration of a light signal from the celestial sphere surface to the centre of Earth – for Earth it is 58.13 days (Table 1, row 4). This value is the reciprocal of the Gaussian gravitational constant confirming planetary-specific finite-radius celestial spheres and aberration angles according to Kepler's third law of planetary motion and Newton's law of gravitation. In terms of β, Kepler's law becomes

$$r_k = \frac{4\pi^2 K}{c^2} = \beta^2 R, \tag{19}$$

where K is Kepler's constant and the radius r_k is half the value of the Schwarzschild radius. Newton's law becomes

$$r_k = \frac{GM}{c^2} = \beta^2 R \tag{20}$$

with G the gravitational constant, and M the mass of the solar system. The value of r_k equals 1476.24711 m. The relation between the Kepler radius (or half the Schwarzschild radius), the parallactic aberration angle and the one radian of an orbit may lead to the concept of discrete radii of the celestial sphere with the vanishing of Earth's imaginary orbit in perspective to infinity, depending on the optical resolution of detection.

6 Multiple discrete stellar aberrations

The one radian of a circle is the equivalent of a phase shift of 1 rad between planetary motion and the arrival time of light from the surface of the celestial sphere, which can be understood from considering the planetary orbit and its imaginary orbit at distance C. Both orbits are in phase, but a light signal requires time to arrive and during the delay the planet travels a curved distance equal to its orbital radius R. This phase difference of 1 rad, noting its association with the radius $C = \beta^{-1}R$ may suggest additional radii $\beta^{-n}R$, because the phase-shift will be precisely 1 rad under these conditions. For

Earth, when n equals 2, the flight time of light is a little over $1\,602$ light-years, i.e. equivalent to $1\,602$ orbital revolutions. The wobbling (hula hoop) of the celestial sphere traces out the planetary orbit at $\beta^{-1}R$ and $\beta^{-2}R$, and thus, light from $\sim$58 days ago and from $\sim1\,602$ years ago are simultaneously observed along the same celestial sphere spindle and in-phase. The vanishing of imaginary planetary orbits at discrete distances in perspective and by virtue of the visibility of the stars by vanishing stellar aberrations (20.49 arcsec, 0.002 arcsec, ...) in perspective is a powerful mechanism to observe depth. Instead of having a celestial sphere with an infinite radius to measure parallactic displacement of stars in the opposite direction of motion, a finite-radius celestial sphere causes negative parallax of all the stars, not some (Fig. 2). Each depth marker on a spindle defines the coordinate at a distance and motion of the observer perpendicular to the line-of-sight does not significantly alter the coordinates at the depth markers faraway with respect to the line-of-sight, in contrast to depth markers nearby. Multiple markers along a spindle and vanishing parallel lines in perspective provide depth perception because the line-of-sight cross spindles when the coordinate system is in motion. The multitude of discrete radii for a given celestial sphere and a fixed line-of-sight along a spindle, i.e. when the telescope is not adjusted, will scan a circular area of the firmament creating a radial field of view of 90° ($\beta^0 = 1$, cf. Fig. 2A), a radial field of view of 20.49 seconds of an arc (β^1, cf. Fig. 2C), a radial field of view of 0.002 seconds of an arc (β^2), a radial field of view of 0.2 microseconds of an arc (β^3), and so on, centered on an imaginary Sun in an anti-clockwise fashion. Large scale rotations suggesting a cosmic web have been reported recently. For example, galaxy rotation appeared to be considerably coherent with the average line-of-sight motion of neighbors at far distances (1–6 Mpc). These rotations are counterclockwise and have a mean velocity at $\sim$30.6 km/s [4], which resembles Earth's orbital velocity. The values reported are consistent
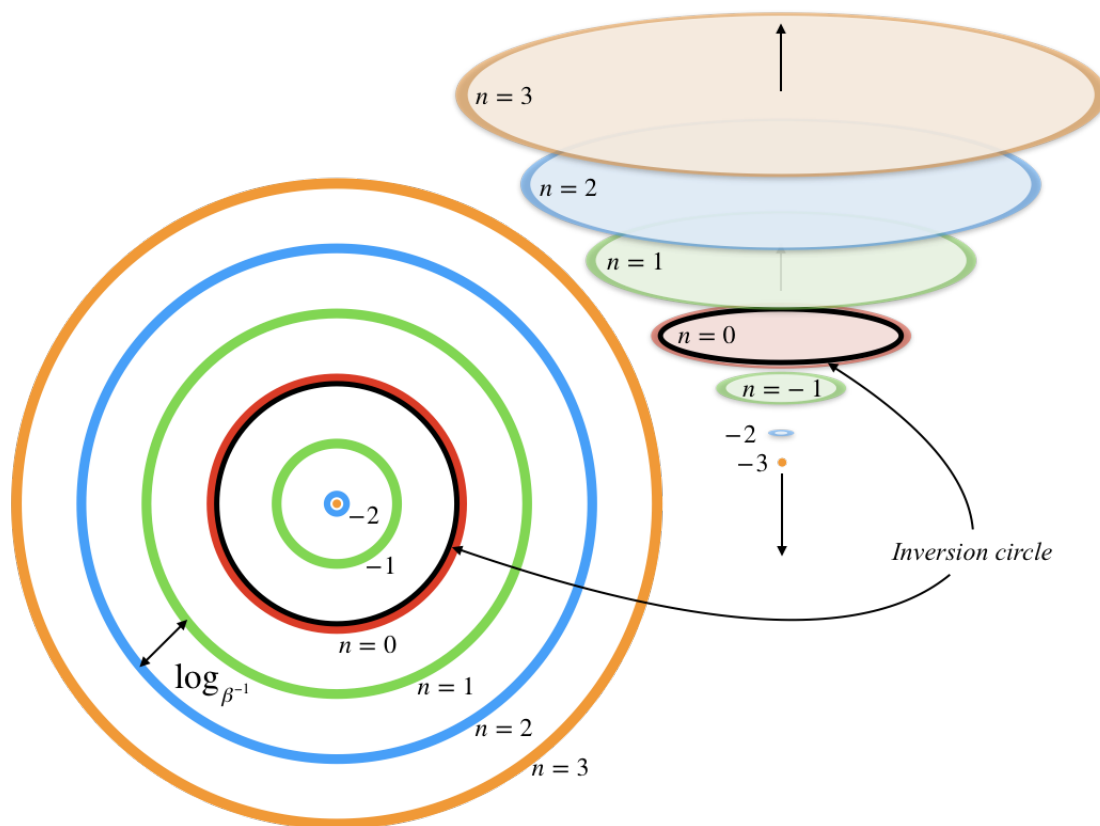
Fig. 5: **Inversion Circle.** The logarithmic scale with base β^{-1} separates the circles (left and right panels). The inversion circle (black) is identical to the planetary orbit (red) when $n = 0$ (24). The green circles represent a circle with radius R/β equal to C when $n = 1$, the principal celestial sphere radius as described in this paper, and its inverse with the inverse radius βR when $n = -1$. Likewise, the blue circle with $n = 2$ is the secondary radius of the celestial sphere (cf. (25)) and its inverse, blue circle with $n = -2$, is the Kepler radius r_k (19–20). The 3-dimensional view (right panel), when viewed from the top, shows the discrete vanishing orbit of Earth in perspective and when viewed from the bottom indicates the stepwise increase of the radius of the celestial sphere.

with a radial field of view equivalent to β^3. Stellar aberration correction in the context of an infinite-radius celestial sphere overcorrects the position of the stars. It causes the fixed stars to have positive parallax and inadvertently make them nearby stars. Instead, abandoning corrective measures and recognizing the finite-radius celestial sphere, stars or nebulae exhibiting positive parallax above 0.002 arcsec are within 1 602 light-years of Earth, and those with less than 0.002 arcsec but above 0.2 μarcsec are within 16 million light-years (4.9 Mpc) from Earth.

7 Solar barycentre precession

With respect to Earth, motion of the other planets add additional aberration of the fixed stars by the wobbling Sun because the solar system barycentre is composed by the individual barycenters for each planet[*]. Sun and planet share a common celestial sphere because the angular velocity, centered on the barycentre of the Sun and planet orbits, is iden-

_____
[*]en.wikipedia.org/wiki/Barycentric_coordinates_(astronomy)

tical, but the orbital velocities of Sun and planet are different and so are the subtending angles that define stellar aberration when viewed from the Sun and planet, respectively. Table 1 (rows 3–6) tabulates specific planetary-based values of Sun's offset to and orbital velocity around the individual barycentre. Major contributors to affect the common barycentre are Jupiter, Saturn, Neptune and Uranus in that order. In return, orbits of planetary celestial spheres change with the periods of the outer planets adding offsets to stellar aberration. Focusing on the effect of Jupiter has on each of the planetary celestial spheres, the torque produced by Jupiter on the Sun adds a wobble with a period of 11.9 Earth-years. The stellar aberration from the Sun orbit around the Sun-Jupiter barycentre equals a subtending angle of 0.00858 seconds of an arc (Table 1, row 7) and becomes an independent component of stellar aberration as observed on Earth. Saturn, Neptune, Uranus contribute significantly and affect the solar system barycentre radius, giving rise to a precession of the celestial sphere coordinates, or in other words, the counterclockwise precession of the solar system barycentre will be "written" at each and

every point of the coordinate system. The barycentre-induced wobble of the Sun might explain S-02's motion centered on Sagittarius A* that has currently a period of about 15 years, has an inclination similar to the ecliptic with respect to the galactic centre and matches the Sun's current orbit and period around the solar system barycentre. The year-by-year (Earth year) subtending angle of the apparent orbit of S-02 as seen from Earth matches the 0.00858 seconds of an arc (Table 1, row 7). Additional contributions are caused by Saturn and Uranus because they are currently located relatively close to Jupiter's position. S-02 is known to rotate in a clockwise direction and is consistent with the notion that Jupiter lags Earth's motion and Earth is the reference against which Jupiter will have the clockwise direction as has the Sun. We note that there are other stars revolving Sagittarius A* with significant longer periods and different inclinations, not associated with the solar barycentre or the ecliptic; this does not take away a possible explanation for S-02's motion.

8 Geometric inversions about an inversion circle

The tracing out of multiple imaginary orbits at discrete distances (Fig. 5), according to

$$C_n = \beta^{-n} R, \qquad (21)$$

suggests also considerations when $n \leq 0$. If $n = 0$ it follows that the celestial sphere radius C_0 equals the orbital radius R and (7) becomes (cf. (6))

$$OQ = R \cos \theta \pm R \cos \theta, \qquad (22)$$

because $\theta = \phi$. This scenario is the equivalent of the fictitious annual circulation (negative parallax) of the Sun through the zodiac. When $n < 0$, and by generalizing (19), (cf. (21)) we get

$$r_k = \beta^{2+n} C_n \qquad (23)$$

celestial sphere radii less than the orbital radius. Since the vorticity, i.e. the velocity field, is determined by the angular velocity of the orbit, the tangential velocity of the celestial sphere coordinates will be less than the orbital velocity. Defining u_C as the tangential velocity of the coordinates and Ω as the angular velocity of the planetary orbit and utilizing (21) for $n \in N$, and defining u'_C when n is negative, we get

$$\begin{aligned} u_C &= \Omega \times C_n = \Omega \times \beta^{-n} R = \beta^{-n} v \\ u'_C &= \Omega \times C'_n = \Omega \times \beta^n R = \beta^n v. \end{aligned} \qquad (24)$$

These equations suggest geometric inversion of points on the circle C_n to their inverse points on circle C'_n with respect to an inversion circle with inversion centre E and inversion radius R. The points on circles C_n and C'_n obey $C_n C'_n = R^2$, the defining feature of circle inversion. Furthermore, the radius $C'_2 = \beta^2 R$, identical to the Kepler radius r_k (19), has its inverse

$$r'_k = \beta^{-2} R \qquad (25)$$

as defined by C_2. The tangential velocities of points and their inverse points (the coordinates) obey

$$u_C u'_C = v^2 \Leftrightarrow \frac{u'_C}{v} = \frac{v}{u_C}. \qquad (26)$$

The Sun's position O and its inverse O', where $O' = O$, are located on the inversion circle. The line-of-sight from the inversion centre E (Fig. 4) to point Q on C_1 harbors Q' on C'_1 (5). In other words, when n is positive, the fictitious orbit is located at the radial distance C_n; when n is negative, C'_n is the radius of the vanishing fictitious orbit in perspective (Fig. 5).

9 Velocity field of the orbit of the Sun

Diurnal and annual motion of the heavens led to paradigm reversals, leading to the first and second motions of the Earth. The third motion of Earth, now known as the axial precession, was, in ancient and medieval times, ascribed to the precession of the equinoxes, a westward motion of the equinoxes along the ecliptic relative to the fixed stars in a cycle of 25 776 years. Precession affected all fixed stars as well as the apparent position of the Sun relative to the backdrop of the stars. The heavens slowly regress a full 360° through the zodiac at the rate of 50.3 seconds of arc per year[*]. Also, other ancient astrologers discovered that the equinoxes "trepidated", particularly along an arc of 46°40' [6], i.e. twice the obliquity of the equinoxes, in one direction and a return to the starting point, resembling how stellar aberration was discovered [2]. The precession and the trepidation appear to be two aspects of the same to-be-proposed frame dependent circulation (Fig. 6), which contrasts Newton's axial precession involving gravitational forces of the Sun and the Moon. Envisioning the Sun orbiting a centre counterclockwise with a period of 25 776 years, with the axis of the Earth in a fixed position (Fig. 6) and noting (18), the radius of the celestial sphere of the Sun becomes 4 102 light-years. Because the equinoctial aberration of the stars is 23.4°, β equals 0.3971, we get an orbital velocity of 119 062 km/s. The radius of the orbit (cf. (21)) is 491.5 pc, or 1 602 light-years and Kepler's radius (19) equals 75.0 pc, which translates to 1.57×10^{15} solar masses at the centre of rotation, vastly exceeding (by 4 magnitudes) current estimates of the Milky Way. The inclination of the zodiacal plane with respect to the invariable plane of the Milky Way galaxy may suggest that the centre of the precession of the equinoxes is not Sagittarius A*. While these values are staggering, concerning or exciting, there may be truth from ancient recordings.

10 Doppler shift measurements from a non-inertial reference frame

Spectroscopic measurement of electromagnetic radiation requires knowledge of Earth's motion, which includes not only the first and second motions, but also the third (Section 9) and

---

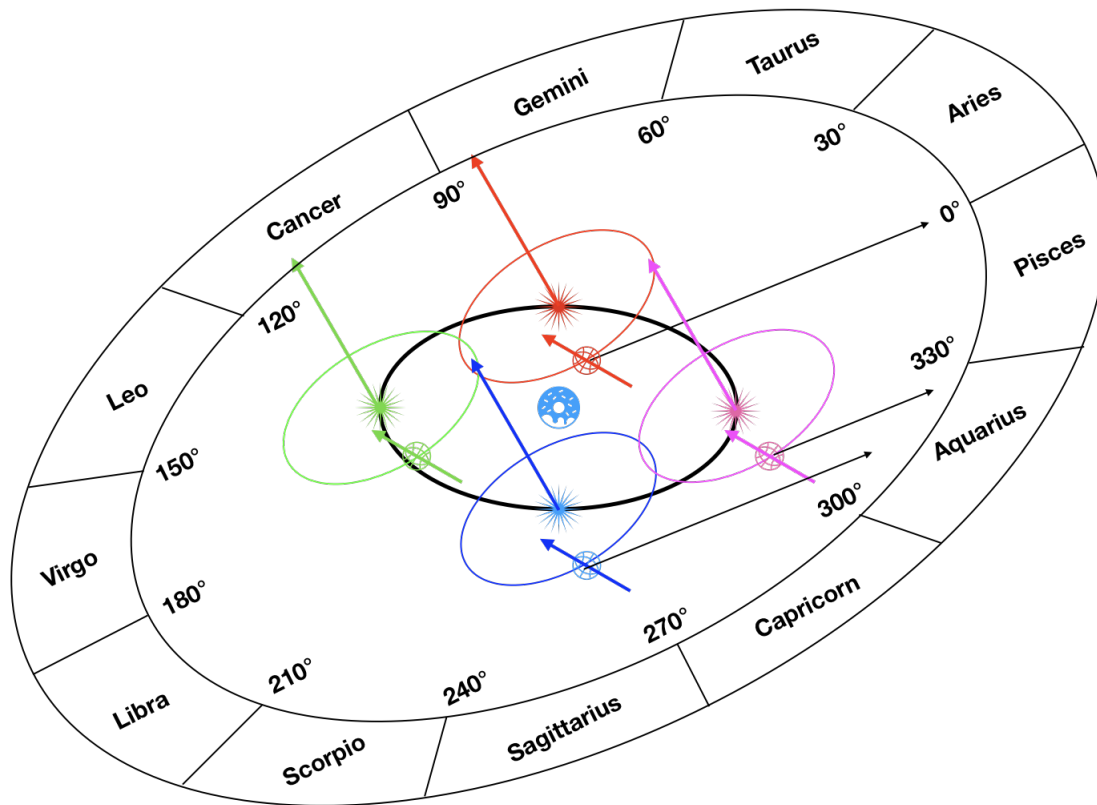[*]en.wikipedia.org/wiki/Axial_precession

Fig. 6: **Stellar aberration by the precession of the equinoxes.** The orbit of Earth around the Sun is shown at four different positions of the Sun in its orbit about the equinoctial center shown as a blue-colored donut shaped symbol. The color coded arrows centered on Earth depicts the north celestial pole of Earth's celestial sphere. The color coded arrows centered on the Sun defines the north celestial pole of the celestial sphere of the Sun. It is assumed that the orbit of the Sun has an obliquity of 30°. Sun's orbit causes stellar aberration with an angular distance of 23.4°, causing equinoctial precession aberration at the vernal equinox. The line-of-sight in the direction of the first point of Aries (0°) is shown by a black arrow. Equinoctial aberration occurs in all directions of the line-of-sight similar to the oscillatory motion of the annual stellar aberration caused by Earth's motion around the Sun.

higher motions. The transformations (13), according to matrix $\mathbf{L}_\odot$ (16), transform the arguments of a sinusoidal wave,

$$\omega t' = k \mathbf{r}', \qquad (27)$$

where ω is the angular frequency and k is the angular wavenumber of the waveform, to

$$\omega \sec \phi \left(t - \mathbf{r} c^{-1} \sin \phi \right) = k \sec \phi \left(\mathbf{r} - ct \sin \phi \right). \qquad (28)$$

Because $\omega k^{-1} = c$, the above identity after rearranging becomes

$$\omega t \sec \phi \left(1 + \sin \phi \right) = k \mathbf{r} \sec \phi \left(1 + \sin \phi \right). \qquad (29)$$

We note that the velocity u of the wave is not affected because

$$u = \frac{\mathbf{r}}{t} = \frac{\omega \sec \phi \left(1 + \sin \phi \right)}{k \sec \phi \left(1 + \sin \phi \right)} = c, \qquad (30)$$

which is the defining feature of the Lorentz matrix. The transformation changes the waveform by $\sec \phi (1 + \sin \phi)$ in the frequency and wavenumber domains with $\sin \phi$ representing the

direction cosine, i.e. the line-of-sight. This Doppler effect, alternatively expressed in terms of $\sin \theta$ (7–9) along the line-of-sight relative to the position of the planet in its orbit, yields a shift z by

$$z + 1 = \sec \phi (1 + \sin \phi) = \sqrt{\frac{1 + \beta \sin \theta}{1 - \beta \sin \theta}}. \qquad (31)$$

Eq. (31) is applicable to any circular motion obeying Kepler's law, such as the Global Position System that sends radio signals. Receivers on Earth will detect a changing Doppler shift depending on the line-of-sight ϕ. Another example tied to (31) is the meaning of a gravitational redshift, which is the equivalent of traversing the Kepler circulation encountering an ever decreasing vorticity with increasing radial distance (19). The cosmological redshift, representing the expansion of the coordinate system, allows speeds greater than the speed of light, where v represents the speed of expansion, not of motion. However, when the ratio of frequencies is equated with

(peculiar) motion v/c,

$$z + 1 = \frac{\omega_{receiver}}{\omega_{source}} = \frac{v}{c} + 1, \qquad (32)$$

and noting the last reported value $z = 11.1$ from galaxy GN-z11, the matrix identities (16,17) and their interpretation (Section 4) do not indicate incompatibility with the Doppler effect. Hence, the expanding universe is likely to be illusionary. In the context of Earth and its celestial sphere being subject to a multitude of circular motions and considering (31) that has two unknowns, θ and β, they may indicate novel periodic changes of the Doppler shift. It is possible that a multitude of orbits would temporally increase Earth's nominal velocity to go beyond the speed of light. The detection of frequency shifts of the order of GN-z11, using $\mathbf{L}_\odot$ (16), may indicate this and if assuming the Milky Way has an orbital velocity close to the speed of light and assuming $\sin\theta = 1$, we have $\beta = 0.984$ for $z = 11.1$. So (32) becomes (cf. (31)),

$$\frac{\omega_r^2}{\omega_s^2} = \frac{1 + \beta\sin\theta}{1 - \beta\sin\theta} \Rightarrow \beta\sin\theta = \frac{\omega_r^2 - \omega_s^2}{\omega_r^2 + \omega_s^2}. \qquad (33)$$

Stars and nebulae beyond the Milky Way may sometimes shine very bright or become very dim at the firmament if Earth obtains a nominal speed equal to the speed of light. Beyond the speed of light the square root of (31) produces an imaginary, inverse result. Geometric inversion (cf. (24)) means passing the luminal barrier, such that

$$\frac{u}{c} = \frac{c}{u'} \Rightarrow uu' = c^2, \qquad (34)$$

where $u \le c$ and $u' \ge c$. We get

$$\frac{\omega_r^2}{\omega_s^2} = \frac{u' + c\sin\theta}{u' - c\sin\theta} = \frac{c^2/u + c\sin\theta}{c^2/u - c\sin\theta} = \frac{c + u\sin\theta}{c - u\sin\theta}, \qquad (35)$$

which is identical to (31). The waveform emanating from a body with velocity u' is not different from a waveform emanating from a body with velocity u and the velocity of our galaxy with respect to the speed of light might be either u or u'. We might not know, but acknowledging frame-dependent induced negative parallaxes may shed further insight in what the universe looks like.

11   Conclusions

Illusion, paradox and true-to-reality phenomena are intertwined in our current worldview, governed by an infinite-radius celestial sphere and merged with the theory of relativity that suggests that space and time are not absolute. The human perception of the third dimension of the universe, be it relativistic or classical in nature, suffers from reversed perspective. The transformations to detect the location of a point by light is governed by the set of the Lorentz $\mathbf{L}$, the Galilean $\mathbf{G}$ and the light retardation $\mathbf{T}$ matrices. Vice versa, light signals

from objects are transformed by premultiplying the Lorentz matrix with the inverse of the light retardation matrix to obtain the Galilean transform. The finite-radius celestial sphere, providing true-to-reality perception, changes the sign of the direction of stellar aberration and therefore parallax of the coordinate system. This recognition may explain co-rotating satellite systems such as a large-scale structure of the universe or the cosmic web. Another large structure of the universe is envisioned based on a fixed Earth axis and a solar orbit with a 1 602 light-year radius about an intragalactic centre causing the precession of the equinoxes with an equinoctial stellar aberration of $23.4°$. The infinite-radius celestial sphere is a relic of ancient times when Earth was considered the centre of the universe. When the second motion of Earth became main stream physics, it should have been accompanied with a finite-radius celestial sphere. It did not because stellar aberration was not discovered until 300 years ago.

References

1. Purves D., Andrews T. J. The perception of transparent three-dimensional objects. *Proc. Natl. Acad. Sci. USA*, 1997, v. 94, 6517–6522.

2. Bradley J. A Letter from the Reverend Mr. James Bradley Savilian Professor of Astronomy at Oxford, and F.R.S. to Dr. Edmond Halley Astronom. Reg. &c. Giving an Account of a New Discovered Motion of the Fix'd Stars. *Philos. Trans. R. Soc.*, 1729, v. 35, 637–661.

3. Einstein A. Zur Elektrodynamik bewegter Körper. *Ann. Phys.*, 1905, v. 322, 891–921.

4. Lee J. H., Pak M., Song H., Lee H. R., Kim S., Jeong H. Mysterious Coherence in Several-megaparsec Scales between Galaxy Rotation and Neighbor Motion. *Astrophys. J.*, 2019, v. 884, 104.

5. Newcomb S. A Compendium of Spherical Astronomy. Macmillan Co., New York,,, 1906 p. 90.

6. Pingree D. Precession and Trepidation in Indian Astronomy before A.D. 1200. *J. Hist. Astron.*, 1972, v. 3, 27–35.

Propagation of a Particle in Discrete Time

Young Joo Noh

yjnoh777@gmail.com

In the concept of discrete time, we can guess the causal delay. A new analysis of causal delays in the dynamics can provide views of two different worlds: type 1 and type 2. In the case of a free particle, the evolution operator for each of them was obtained and analyzed. As a result, type 1 particle could be interpreted as ordinary matter that satisfies existing relativistic quantum mechanics. Type 2 particle is outside the quantum mechanics category, but has some interesting physical properties. Type 2 particle acts on gravity in the same way as ordinary matter, and does not interact with the U(1) gauge field, and considering its energy density value, it can be interpreted as dark matter.

1 Introduction

The dynamical system aims to find dynamic variables that change over time, which is the process of solving the equations of motion. The structure of the equation of motion combines the amount of change of the dynamic variables with time and the cause of the change.

As an example, let's take a look at Newton's laws of motion. Newton expressed his second law as follows:

> *Change of motion is proportional to impressed motive force and is in the same direction as the impressed force.*

The equation of motion is as follows.

$$\lim_{\Delta t \to 0} \frac{\Delta \vec{p}}{\Delta t} = \vec{F}(t) \, .$$

Applying the cause-effect category to Newton's law of motion mentioned above, force is the cause and momentum changes are the effects.

However, the point to note here is the time difference between the moment t when the cause force is applied and $t + \Delta t$ which is the moment when the resultant momentum change appears. Naturally, in continuous space and time, this time difference is infinitely small, so the cause and effect are "simultaneous". Let's call this simultaneity *infinitesimally different simultaneity*. This infinitesimally different simultaneity is assumed in all dynamic systems based on continuous space and time: Newtonian mechanics, Lagrangian mechanics, Hamiltonian mechanics, Quantum mechanics, etc.

By the way, this infinitesimally different simultaneity is two different points, unlike *true simultaneity* which is identical. Because if they are the same, then at any moment an object has to have both momentum before the cause and momentum after the cause. The distinction between two points in infinitesimally different simultaneity in continuous space and time is meaningless, but in discrete time, there is a minimum value Δt_d for time change and two points for cause and effect, resulting in a delay of time Δt_d between cause and effect.

The delay between cause and effect will of course affect the description of the dynamics, which requires an evolution operator for a particle in discrete time. There are two types of results in this process, one that is consistent with existing relativistic quantum mechanics and another that is entirely new.

2 Definitions

2.1 Cause-effect vectors

Considering the causal delay, we cannot define the "real state" at one moment, as in quantum mechanics, and define the "real state" within the minimum time Δt_d. The existing quantum mechanical state with 4-momentum p_μ at a point x^α in spacetime can be called $\phi_p(x^\alpha)$ which is caused by $x^\alpha - \Delta x^\alpha$ or $x^\alpha + \Delta x^\alpha$ due to the causal delay. Where Δx^α is a timelike 4-vector, meaning cause-effect delation in space-time. The time component of Δx^α is an amount representing the cause-effect time delation Δt_d and the spatial component represents the distance the object moved during the time delay.[*]

Therefore, the definition of the "real state" in discrete time must be made by combining coordinate values with $\phi_p(x^\alpha)$. So there are two definitions, past-future cause-effect vector and future-past cause-effect vector. Where $\phi_p(x^\alpha)$ is tentatively scalar.

past-future cause-effect vector : $x^\mu \phi_p(x^\alpha + \Delta x^\alpha)$,

future-past cause-effect vector : $(x^\mu + \Delta x^\mu)\phi_p(x^\alpha)$.

2.2 Difference of cause-effect vectors

Since there are two states between Δx^α as discussed above, by combining them, the state change can be of two types:

$$\text{type 1} : (x^\mu + \Delta x^\mu)\phi_p(x^\alpha) - x^\mu \phi_{p'}(x^\alpha + \Delta x^\alpha) \, . \tag{1}$$

$$\text{type 2} : x^\mu \phi_{p'}(x^\alpha + \Delta x'^\alpha) - (x^\mu - \Delta x^\mu)\phi_p(x^\alpha) \, . \tag{2}$$

---

[*]In this paper, unlike time, distance in space does not assume its minimum value. The discreteness of space is a controversial topic and has nothing to do with the content of this paper.

The two types are shown schematically in Fig. 1.* As shown in Fig. 1, type 1 is the difference between future-past cause-effect vector and past-future cause-effect vector, and type 2 consists only of the difference between past-future cause-effect vectors.
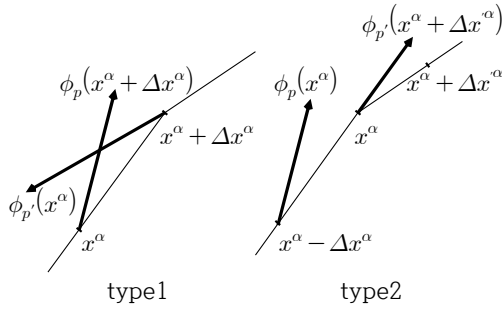

Fig. 1: Schematics of type 1 and type 2

3　Calculations of difference of cause-effect vectors for a free particle

Assumption 1 : In type 1, the state value $\phi_p(x^\alpha)$ at point x^α has the same magnitude of contribution at $x^\alpha - \Delta x^\alpha$ and $x^\alpha + \Delta x^\alpha$.
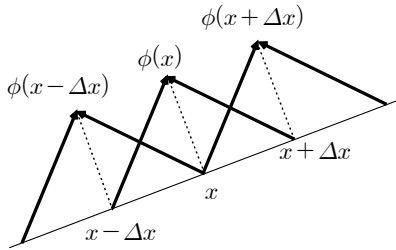
Fig. 2: Contributions to each ϕ

In Fig. 2, $\phi(x)$ is a mixture of contributions from $x - \Delta x$ and $x + \Delta x$. The same applies to the other ϕ's. This can be written in the following way. p is omitted because it is the same.

$$\phi(x) = \phi_{x-\Delta x}(x) + \phi_{x+\Delta x}(x) . \tag{3}$$

This contribution in space-time is shown in Fig. 3.

As shown in the Fig. 3, the two contributions to $\phi(x)$ at x^μ will act in opposite directions on the tangent of the dotted world line. Thus both contributions will be offset. The same is true for vectors. That is, scalars and vectors cannot describe type 1.

---

*In discrete time, the trajectory in space-time cannot be a solid line, only a jump from point to point. The solid trajectory in Fig. 1 and 2 is just for readability.
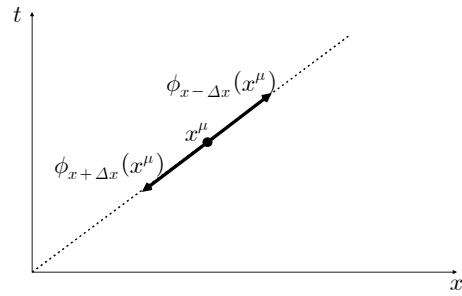
Fig. 3: two contributions to $\phi(x)$

What about the spinor? In Fig. 3, the two spinors are also the same magnitude and in opposite directions, but the sum is not zero. Because in spinor space the two spinors are orthogonal. Two orthogonal spinors correspond to $\binom{1}{0}$ and $\binom{0}{1}$, respectively, which correspond to spin $\frac{1}{2}$. Thus, only spin $\frac{1}{2}$ spinors can describe type 1.

Now, if the spinor is constant, the difference of cause-effect vectors for a 2-component spinor $\Psi_a(x)$ is defined as follows.

$$\text{type 1} : (x^\mu + \Delta x^\mu) \Psi_a(x) - x^\mu \Psi_a(x + \Delta x) . \tag{4}$$

In the case of a free particle, the difference of cause-effect vectors for type 2 is also shown.

$$\text{type 2} : x^\mu \phi(x + \Delta x) - (x^\mu - \Delta x^\mu) \phi(x) . \tag{5}$$

Assumption 2 : Ψ and ϕ are analytic functions.

But in reality it is discontinuous and it is difficult to figure this out. This assumption approximates discontinuous Ψ and ϕ as C^∞ functions, which means looking at the dynamical point of view that we are familiar with.

3.1　Type 1

Let's express the spinor function $\Psi_a(x)$ as a spinor part and a scalar part depending on the coordinates as follows.

$$\Psi_a(x) = u_a \phi(x) . \tag{6}$$

Then, (4) is as follows.

$$u_a \{(x^\mu + \Delta x^\mu) \phi(x) - x^\mu \phi(x + \Delta x)\} . \tag{7}$$

So we only need to calculate the part for scalar.

$$(x^\mu + \Delta x^\mu) \phi(x^\alpha) - x^\mu \phi(x^\alpha + \Delta x^\alpha)$$

$$= (x^\mu + \Delta x^\mu) \phi(x^\alpha) - x^\mu \sum_{n=0}^{\infty} \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n \phi(x^\alpha)$$

$$= \Delta x^\lambda \left\{\delta_\lambda^\mu \phi(x^\alpha) - x^\mu \frac{\partial \phi(x^\alpha)}{\partial x^\lambda}\right\} -$$

$$- x^\mu \sum_{n=2}^{\infty} \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n \phi(x^\alpha) .$$

For $n \geq 2$

$$x^\mu \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n = \left[x^\mu, \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n \right]_{\text{commutation}} .$$

Thus

$$(x^\mu + \Delta x^\mu) \phi(x^\alpha) - x^\mu \phi(x^\alpha + \Delta x^\alpha)$$

$$= -\left[x^\mu, \Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right] \phi(x) -$$

$$- \left[x^\mu, \sum_{n=2}^\infty \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n \right] \phi(x)$$

$$= -\left[x^\mu, \sum_{n=1}^\infty \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n \right] \phi(x)$$

$$= -\left[x^\mu, \exp\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right) - 1 \right] \phi(x)$$

$$= -\left[x^\mu, \exp\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right) \right] \phi(x) .$$

For the progress of the calculation, we define the 4-momentum operator P_λ, and commutation relation of x^μ and P_λ as follows.

$$P_\lambda \equiv i\hbar \frac{\partial}{\partial x^\lambda}$$

$$[x^\mu, P_\lambda] \equiv -i\hbar \delta_\lambda^\mu \tag{9}$$

and metric $\eta_{\alpha\beta} = \text{diag}\begin{bmatrix} 1 & -1 & -1 & -1 \end{bmatrix}$.

Using the following (10), the final result is as shown in (11).

$$[x_i, F(P_i)] = i\hbar \frac{dF}{dP_i}$$

$$[x_0, F(P_0)] = -i\hbar \frac{dF}{dP_0} . \tag{10}$$

$$(x^\mu + \Delta x^\mu) \phi(x^\alpha) - x^\mu \phi(x^\alpha + \Delta x^\alpha)$$

$$= \Delta x^\mu \exp\left(-\frac{i}{\hbar} \Delta x^\alpha P_\alpha \right) \phi(x) . \tag{11}$$

Therefore, the equation for spinor function $\Psi_a(x)$ is

$$(x^\mu + \Delta x^\mu) \Psi_a(x) - x^\mu \Psi_a(x + \Delta x)$$

$$= \Delta x^\mu \exp\left(-\frac{i}{\hbar} \Delta x^\alpha P_\alpha \right) \Psi_a(x) . \tag{12}$$

3.2 Type 2

After a similar calculation process as in type 1, the equation that corresponds to (12) is (16).

$$x^\mu \phi(x + \Delta x) - (x^\mu - \Delta x^\mu) \phi(x)$$

$$= x^\mu \sum_{n=0}^\infty \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n \phi(x) - (x^\mu - \Delta x^\mu) \phi(x)$$

$$= \Delta x^\alpha \left(x^\mu \frac{\partial}{\partial x^\alpha} + \delta_\alpha^\mu \right) \phi(x) + x^\mu \sum_{n=2}^\infty \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n \phi(x) .$$

For $n \geq 2$

$$x^\mu \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n = \left\{ x^\mu, \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n \right\}_{\text{anticommutation}} .$$

$$x^\mu \phi(x + \Delta x) - (x^\mu - \Delta x^\mu) \phi(x)$$

$$= \left\{ x^\mu, \Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right\} \phi(x) + \left\{ x^\mu, \sum_{n=2}^\infty \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n \right\} \phi(x)$$

$$= \left\{ x^\mu, \sum_{n=1}^\infty \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right)^n \right\} \phi(x) \tag{13}$$

$$= \left\{ x^\mu, \exp\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha} \right) - 1 \right\} \phi(x) .$$

Note that unlike type 1, anticommutation occurs. Therefore, for calculation, we need to define the anticommutation relation of 4-vector x and P as below.

$$P_\lambda \equiv i\hbar \frac{\partial}{\partial x^\lambda}$$

$$\{x^\mu, P_\lambda\} \equiv i\hbar \delta_\lambda^\mu . \tag{14}$$

And using the following (15), the final equation corresponding to (12) of type 1 is the following (16).

$$\{x_i, G(P_i)\} = -i\hbar \frac{dG}{dP_i}$$

$$\{x_0, G(P_0)\} = i\hbar \frac{dG}{dP_0} . \tag{15}$$

$$x^\mu \phi(x + \Delta x) - (x^\mu - \Delta x^\mu) \phi(x)$$

$$= \left(\Delta x^\mu \exp\left(-\frac{i}{\hbar} \Delta x^\alpha P_\alpha \right) - 2x^\mu \right) \phi(x) . \tag{16}$$

To understand the meaning of the right sides of (12) and (16) for type 1 and type 2, we first briefly review the time evolution operator in quantum mechanics in the next chapter. In a similar manner, in space-time, the right side of (12) will be defined as the evolution operator of the type 1 particle and the right side of (16) will be defined as the evolution operator of the type 2 particle.

4 Evolution operator

In quantum mechanics, the time evolution operator U from state $|\alpha, t_0\rangle$ to $|\alpha, t_0; t\rangle$ is defined as follows.

$$|\alpha, t_0; t\rangle = U(t, t_0) |\alpha, t_0\rangle .$$

Determining this time evolution operator is equivalent to determining the equation of motion for the state of the system. If the time evolution operator is as follows, this operator satisfies the Schrodinger equation.

$$U(t, t_0) = \exp\left(-\frac{iH(t - t_0)}{\hbar} \right)$$

$$i\hbar \frac{\partial U(t, t_0)}{\partial t} = HU(t, t_0)$$

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) |\alpha, t_0\rangle = HU(t, t_0) |\alpha, t_0\rangle .$$

That is, the Schrodinger equation for the state is established as below.

$$i\hbar \frac{\partial}{\partial t} \mid \alpha, t_0; t \rangle = H \mid \alpha, t_0; t \rangle .$$

Now let's discuss type 1 and type 2. Type 1 and type 2 discussed here are free particles, so we can apply the concept of evolution operator in quantum mechanics.

4.1 Type 1

$$(x^\mu + \Delta x^\mu) \Psi_a (x) - x^\mu \Psi_a (x + \Delta x)$$

$$= \Delta x^\mu \exp\left(-\frac{i}{\hbar} \Delta x^\alpha \hat{P}_\alpha\right) \Psi_a (x) \qquad (17)$$

$$\equiv \Delta x^\mu U (\Delta x) \Psi_a (x) .$$

Suppose $U (\Delta x)$ in (17) is an evolution operator by Δx in space-time. Then

$$\Psi_a^p (\Delta x) \equiv U (\Delta x) \Psi_a^p (0)$$
$$= \exp\left(-\frac{i}{\hbar} \Delta x \cdot \hat{P}\right) \Psi_a^p (0) = \exp\left(-\frac{i}{\hbar} \Delta x \cdot p\right) \Psi_a^p (0) . \qquad (18)$$

Successive evolution by $n\Delta x^\alpha = x^\alpha$ can be expressed as below.

$$\exp\left(-\frac{i}{\hbar} \Delta x \cdot \hat{P}\right) \cdots \exp\left(-\frac{i}{\hbar} \Delta x \cdot \hat{P}\right) \Psi_a^p (0)$$

$$= \exp\left(-\frac{i}{\hbar} n\Delta x \cdot p\right) \Psi_a^p (0) = \exp\left(-\frac{i}{\hbar} x \cdot p\right) \Psi_a^p (0)$$

$$= \exp\left(-\frac{i}{\hbar} x \cdot \hat{P}\right) \Psi_a^p (0) = U (x) \Psi_a^p (0) .$$

Therefore, $U (x)$ satisfies the following equation.

$$\Psi_a^p (x) = U (x) \Psi_a^p (0) . \qquad (19)$$

If we apply the Klein-Gordon operator to (19) and use p as constant, we get the following result ($\hbar = 1$).

$$\left(\partial_\mu \partial^\mu + m^2\right) \Psi_a^p (x)$$

$$= \left(\partial_\mu \partial^\mu + m^2\right) e^{-ix \cdot \hat{P}} \Psi_a^p (0)$$

$$= \left(\partial_\mu \partial^\mu + m^2\right) e^{-ix \cdot p} \Psi_a^p (0) \qquad (20)$$

$$= \left(-p_\mu p^\mu + m^2\right) e^{-ix \cdot p} \Psi_a^p (0)$$

$$= 0 .$$

As you can see from (20), $\Psi_a^p (x)$ is the solution of the Klein-Gordon equation. And, as discussed above, $\Psi_a^p (x)$ is spin $\frac{1}{2}$, so $\Psi_a^p (x)$ can be said to be a component of a spinor that satisfies the Dirac equation.

In summary, for a free particle, type 1 can be interpreted as a conventional ordinary matter that satisfies the Dirac equation, and $U (x)$ can be interpreted as an evolution operator. It is also worth noting that type 1 particles, although their beginnings are unusual, are in agreement with existing relativistic quantum mechanics, indicating some of the validity of the causal delay.

4.2 Type 2

$$x^\mu \phi (x + \Delta x) - (x^\mu - \Delta x^\mu) \phi (x)$$

$$= \left(\Delta x^\mu exp\left(-\frac{i}{\hbar} \Delta x^\alpha P_\alpha\right) - 2x^\mu\right) \phi (x) \qquad (21)$$

$$\equiv \Delta x^\mu V \phi (x) .$$

We have dicussed that U of type 1 can be interpreted as an evolution operator. Based on that, we will define V as an evolution operator of type 2. But unlike U, V is not a unitary operator, i.e. type 2 particles are broken in unitarity. Nevertheless, type 2 particles have very interesting physical meanings.

5 Properties of type 2 particle

5.1 $x^\mu \gg \Delta x^\mu$

The evolution operator at large x is

$$\Delta x^\mu V \simeq -2x^\mu \qquad (22)$$

Since V is a linear function of x in (22), the equation that V must satisfy is the second order differential equation, i.e. $\partial_\alpha \partial_\beta V = 0$. Thus, the equation of motion that x must satisfy can be written in covariant form as

$$\frac{d^2 x^\mu}{d\tau^2} = 0 \qquad (23)$$

where τ is the proper time.

Eq. (23) is the classical relativistic equation of motion for a free particle, not a wave equation. This means that in large x there is only motion as a particle and no quantum waves.

More discussion is needed about the above. If we take $\Delta x \to 0$ limit on both sides in (16), it is as follows.

$$x^\mu \left(\phi (x) + \Delta x^\alpha \frac{\partial \phi}{\partial x^\alpha}\right) - (x^\mu - \Delta x^\mu) \phi (x)$$

$$= \Delta x^\mu \left\{\phi (x) - \frac{i}{\hbar} \Delta x^\alpha p_\alpha \phi (x)\right\} - 2x^\mu \phi (x)$$

$$\Delta x^\alpha \frac{\partial \phi (x)}{\partial x^\alpha} = -2\phi (x) .$$

$$\therefore \phi (x) \propto \exp\left(-2 \frac{\Delta x \cdot x}{\Delta x \cdot \Delta x}\right) . \qquad (24)$$

As shown in (24), the particle position is very localized. However, this is the position value in the state where the momentum is determined. In other words, type 2 particle can be determined at the same time the position and momentum, which means that there is no quantum wave phenomenon in the type 2 particle. Although quantum waves do not exist, it has a physical meaning because it satisfies the classical equation of motion.

Eq. (23) holds for an inertial frame in flat spacetime. If a curved spacetime manifold is locally flat at an arbitrary point P, (23) always holds at P because $\frac{\partial}{\partial x^\gamma} g_{\alpha\beta} (P) = 0$. This means

that in locally flat manifolds, type 2 particle undergo free-falling motion with a straight geodesic. That is, type 2 particle is affected by gravity in the same way as ordinary matter.

5.2　$x^\mu \simeq \Delta x^\mu$

The evolution operator in this case is as follows.

$$V \simeq \exp\left(-\frac{i}{\hbar} x \cdot \hat{P}\right) - 2 \,. \tag{25}$$

The first term in (25) is the operator giving the Klein-Gordon equation. The second term, as discussed in 5.1, means acceleration, which is related to mass. Thus, (25) can be seen as an operator that gives an equation that modifies the mass part of the Klein-Gordon equation. If the modified Klein-Gordon equation is set as shown in (26) below, f is obtained as follows ($\hbar = 1$).

$$\left(\partial_\mu \partial^\mu + m^2 f\right) \phi_p(x) = 0 \,. \tag{26}$$

where $\phi_p(x) = V(x)\phi_p(0)$.

$$\left(\partial_\mu \partial^\mu + m^2 f\right)\left(e^{-ix\cdot\hat{P}} - 2\right)\phi_p(0) = 0 \,.$$
$$\therefore f(x) = \frac{e^{-ix\cdot p}}{e^{-ix\cdot p} - 2} \,. \tag{27}$$

In the modified Klein-Gordon equation, the mass term is a complex number.

We will now discuss the internal symmetry of type 2 particles.

The equation that satisfies $\phi(x)$ and $\phi^*(x)$ in (26) is as follows.

$$\left(\partial_\mu \partial^\mu + m^2 f(x)\right) \phi(x) = 0$$
$$\left(\partial_\mu \partial^\mu + m^2 f^*(x)\right) \phi^*(x) = 0 \,. \tag{28}$$

As can be seen from (28), the equations satisfying $\phi(x)$ and $\phi^*(x)$ are different. This means that the type 2 particles do not have antiparticles and, as will be seen later, do not have internal symmetry. To show that the type 2 particles do not have internal symmetry, the Lagrangian density must be determined. However, defining the Lagrangian density implies that type 2 is assumed to be a field only locally, although this is not the case for large x. In addition, the Lagrangian density should be a locally holomorphic complex Lagrangian.

By the way, the Lagrangian density of a normal complex scalar field cannot produce (28). Therefore, some process is required.

First, changing the expression (28) using $f^*(x) = f(-x)$ is as follows.

$$\left(\partial_\mu \partial^\mu + m^2 f(-x)\right) \phi^*(x) = 0 \,.$$

And $x \to -x$ gives

$$\left(\partial_\mu \partial^\mu + m^2 f(x)\right) \phi^*(-x) = 0 \,. \tag{29}$$

In (28) and (29), it can be seen that $\phi(x)$ and $\phi^*(-x)$ satisfy the same equation. Thus Lagrangian density can be written as

$$\mathcal{L} = \partial_\mu \phi^*(-x) \partial^\mu \phi(x) - m^2 f(x) \phi^*(-x) \phi(x) \,. \tag{30}$$

Now consider the following gauge transformations.

$$\phi(x) \to e^{-iq\theta(x)}\phi(x) \,, \quad \phi^*(-x) \to e^{iq\theta(-x)}\phi^*(-x) \,.$$
$$\frac{\delta\phi(x)}{\delta\theta(x)} = -iq\phi(x) \,, \quad \frac{\delta\phi^*(-x)}{\delta\theta(x)} = iq\phi^*(-x)\frac{\delta\theta(-x)}{\delta\theta(x)} \,. \tag{31}$$

According to Noether's theorem, if the action is invariant under gauge transformations, there is a vanishing divergence current, whose value is

$$J^\mu = -\left[\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu \phi(x)\right)}\frac{\delta\phi(x)}{\delta\theta(x)} + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu \phi^*(-x)\right)}\frac{\delta\phi^*(-x)}{\delta\theta(x)}\right]$$
$$= iq\left[\phi(x)\partial^\mu\phi^*(-x) - \phi^*(-x)\partial^\mu\phi(x)\frac{\delta\theta(-x)}{\delta\theta(x)}\right] \,. \tag{32}$$

However, the divergence of current in (32) is not zero. And this means that there is no conserved charge.

Therefore, we can say that Lagrangian density in (30) has no internal symmetry. In other words, type 2 particles cannot construct covariant derivatives that satisfy gauge invariance, which means that type 2 particles do not interact with the U(1) gauge field.

In addition, since type 2 particles lack an intenal symmetry, it means that it is a kind of scalar without components, so SU(2) and SU(3) gauge symmetry cannot be defined. Accordingly, type 2 particles do not have weak interactions and strong interactions as well as electromagnetic interactions, but is connected only by gravity.

5.3　Mass and energy density

Let's first discuss the mass of type 2 particle.

In the modified Klein-Gordon equation, mass is distributed in space-time like a wave, and its distribution is determined by f. Therefore, to find the mass of type 2 particle, we need to find the integral value for the space. Let $t = 0, \vec{p} = (p, 0, 0)$ be for simplicity. Then f is

$$f(x) = \frac{e^{ipx}}{e^{ipx} - 2} \,. \tag{33}$$

In (33), f diverges at $x \to \pm\infty$. This is because the expression (33) holds for $x^\mu \simeq \Delta x^\mu$. Therefore, to find the integral, we need to define a function value at $x \to \pm\infty$. As discussed in section 5.1, type 2 particles do not have a wave function for large x. Consequently, we can set the boundary condition $f \to 0$ at $x \to \pm\infty$. In order for f to converge at $x \to \pm\infty$, we need to modify f in (33). By introducing damping factor ϵ, modified f is presented as below.

$$f_m = \frac{e^{i(p+i\epsilon)|x|}}{e^{i(p+i\epsilon)x} - 2} \,. \tag{34}$$

Therefore, the mass M of type 2 particle can be defined as (35). And m is the mass of ordinary matter that satisfies Klein-Gordon equation.

$$M^2 \equiv m^2 \left| \int_{-\infty}^{\infty} \frac{e^{i(p+i\epsilon)|x|}}{e^{i(p+i\epsilon)x} - 2} dx \right| . \tag{35}$$

In order to calculate the integral value of (35), the following integral of a complex variable must be obtained.

$$\oint dz \frac{e^{i(p+i\epsilon)|z|}}{e^{i(p+i\epsilon)z} - 2} . \tag{36}$$

The poles and residues are

simple pole $\quad z_0 = -\dfrac{(\epsilon + ip)}{\epsilon^2 + p^2} \ln 2 .$

residue at z_0

$$\begin{aligned} a_{-1} &= \frac{e^{i(p+i\epsilon)|z|}}{e^{i(p+i\epsilon)z} - 2} \cdot \left(e^{i(p+i\epsilon)z} - 2 \right)|_{z=z_0} \\ &= 2^{i \frac{(p+i\epsilon)}{\sqrt{p^2+\epsilon^2}}} \\ &= 2^i \quad (\text{for } \epsilon \to 0) . \end{aligned}$$
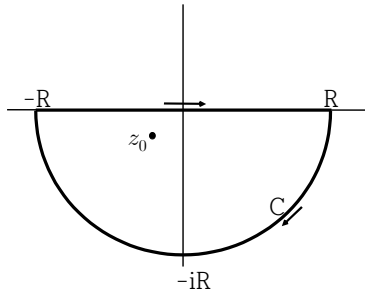
The contour of integration is shown in Fig. 4.

Fig. 4: Contour of integration

$$\begin{aligned} &\oint dz \frac{e^{i(p+i\epsilon)|z|}}{e^{i(p+i\epsilon)z} - 2} \\ &= \lim_{R \to \infty} \int_{-R}^{R} dx \frac{e^{i(p+i\epsilon)|x|}}{e^{i(p+i\epsilon)x} - 2} + \int_{C} dz \frac{e^{i(p+i\epsilon)|z|}}{e^{i(p+i\epsilon)z} - 2} \\ &= 2\pi i a_{-1} = 2\pi i \cdot 2^i . \end{aligned} \tag{37}$$

Since the second integral term in (37) is 0 as $R \to \infty$, the mass value to be obtained is

$$\frac{M^2}{m^2} = \left| 2\pi i \cdot 2^i \right| = 2\pi . \tag{38}$$

Let's discuss the energy density. First, the energy density of type 1, that is, ordinary matter, is as follows in the case of

a complex scalar field.

$$\mathcal{L} = \partial_\mu \phi^*(x) \partial^\mu \phi(x) - m^2 \phi^*(x) \phi(x)$$

$$T^{\mu\lambda} = \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi_i} \partial^\lambda \phi_i - \eta^{\mu\lambda} \mathcal{L} = \partial^\mu \phi^* \partial^\lambda \phi + \partial^\mu \phi \partial^\lambda \phi^* - \eta^{\mu\lambda} \mathcal{L} .$$

Accordingly, the energy density in the case of a free particle is as follows.

$$T^{00} = \left(\vec{p}^2 + m^2 \right) \phi^* \phi .$$

Ignoring the kinetic part:

$$T^{00}_{\text{type 1}} \simeq m^2 |\phi_1|^2 . \tag{39}$$

Lagrangian density and energy momentum tensor of type 2 are as follows.

$$\mathcal{L} = \partial_\mu \phi^*(-x) \partial^\mu \phi(x) - m^2 f(x) \phi^*(-x) \phi(x)$$
$$T^{\mu\lambda} = \partial^\mu \phi^*(-x) \partial^\lambda \phi(x) + \partial^\mu \phi(x) \partial^\lambda \phi^*(-x) - \eta^{\mu\lambda} \mathcal{L} . \tag{40}$$

However, as discussed earlier, the Lagrangian density of type 2 is a complex number, so the energy momentum tensor of the above formula is also a complex number. Therefore, the above energy momentum tensor cannot be applied to the physical system as it is.

This issue is intended to find meaning through the following discussion. As discussed earlier, type 2 has no internal symmetry, so there is no short distance interaction. That is, only long distance interaction (gravity) is possible. However, at far distances, there is no wave property, but only particle properties, so type 2 has only meaning as particles in the long distance interaction. Acting as a particle means that it participates in gravity as a particle having a mass M of the type 2 obtained above. In this case, the behavior of the particles as mass M is equivalent to ordinary matter. Therefore, the energy density of type 2 can be treated as the energy density of the scalar field of mass M. Accordingly, the same process as the energy density of type 1 discussed above is as follows.

$$T^{00}_{\text{type 2}} \simeq M^2 |\phi_2|^2 . \tag{41}$$

Consequently, the energy density ratio of type 1 and type 2 particles is as follows.

$$\frac{T^{00}_{\text{type 2}}}{T^{00}_{\text{type 1}}} \simeq \frac{M^2 |\phi_2|^2}{m^2 |\phi_1|^2} . \tag{42}$$

One thing to note here is that the mass m of ordinary matter compared in the above formula is the mass of the particle as a free particle.

In (42), $|\phi_2|^2 / |\phi_1|^2$ is the ratio of mass-independent amplitudes, so we can make them equal. Accordingly, the following results can be obtained.

$$\frac{T^{00}_{\text{type 2}}}{T^{00}_{\text{type 1}}} \simeq \frac{M^2}{m^2} = 2\pi = \frac{86.3\%}{13.7\%} . \tag{43}$$

Since type 1 and type 2 have the same opportunity for generation in their origin, the number density of the two will be the same. Therefore, the ratio of the above equation is the ratio of the energy density of the total amount of type 1 and type 2 in the universe. The value is within the range of the energy density ratios of dark matter and ordinary matter that are currently estimated.

6 Conclusions

The interpretation of the dynamical system with a new concept of causal delay, which originated from the discrete concept of time, gave us a perspective on two different worlds. For a free particle, type 1 particle can be interpreted as ordinary matter that satisfies existing relativistic quantum mechanics. This type 1 particle can only have spin $\frac{1}{2}$, which can explain why the spin of all fermions observed is $\frac{1}{2}$.

Type 2 particle is a matter of a whole new perspective. This particle does not follow the existing laws of quantum mechanics. Type 2 is only a classical particle that satisfies the theory of relativity at a long distance, and has a property as a kind of field that does not have gauge symmetry at a short distance. So, these type 2 particles act on gravity in the same way as ordinary matter, do not interact with light, and considering their energy density value, it can be interpreted as dark matter.

Type 2 particles do not have any gauge interactions. And there is no antiparticle, including itself, so no annihilation occurs. Therefore, direct or indirect detection based on them is not possible, only indirect verification through gravity. However, given the local nature of type 2, it is not a point-like particle, so self-interaction through collision seems to be possible.

References

1. Whitrow G. J. The Natural Philosophy of Time. Oxford University Press, 1980.

2. Penrose R. The Road to Reality. Jonathan Cape, 2004.

3. Misner C. W., Thorne K. S., Wheeler J. A. Gravitation. W. H. Freeman, 1970.

4. Aldrovandi R., Pereira J. G. Notes for a Course on Classical Fields, IFT, 2008.

5. Sakurai J. J. Modern Quantum Mechanics. The Benjamin/Cummings Publishing, 1985.

6. Bjorken J. D. Drell Sidney D. Relativistic Quantum Mechanics. McGraw-Hill Book, 1964.

7. Gondran M., Kenoufi A., Gondran A. Complex Variational Calculus with Mean of (min,+)-analysis. *Ten. Mat. Apl. Comput.*, 2017, v. 18 (3), 385–403.

8. Bertone G., Hooper D. History of Dark Matter. *Review of Modern Physics*, 2018, v. 90(4), 045002.

9. Bertone G., Hooper D., Silk J. Particle Dark Matter: Evidence, Candidates and Constraints. arXiv:hep-ph/0404175.

10. Sarkar S. Is dark matter self-interacting? *Nature Astronomy*, 2018, v 2, 856–857.

11. Peter A. H. G. Dark Matter. arXiv:astro-ph/1201.3942v1.

12. Steane A. M. An Introduction to Spinors. arXiv:math-ph/1312.3824v1.

13. Crouse D. T. The nature of discrete space-time. arXiv:class-ph/1608.08506v2.

On the Electron Pair, the Single Bond C-C Rotational Energy Barrier and Other Molecular Mechanisms

Omar Yépez

E-mail: yepezoj@gmail.com

To find evidence of the electron pair has proven to be a very difficult task. Bader et al. tried to unsuccessfully find evidence of the electron pair in the topological analysis of the Laplacian of the electron density of molecules. By using electron localization functions, Silvi et al. pointed out where these pairs might be in the molecule and represented them as attractors. Still, to locate the electron pair does not give answers to different molecular mechanisms. For instance, the mechanism of hindered rotation about the carbon-carbon single bond in ethane, which is of great interest and controversy. This phenomenon is not yet explained by Silvi's most advanced molecular model (state of the art). A new alternative uses the relationship between the area of the electron density and the energy of the bond. This approach also provides the electron pair localization. Furthermore, by allowing the magnetic momenta of the bonding electrons to interact, an explanation of the rotational barrier appeared straightforwardly. Also, the model presented in this paper find bonding electrons not found by Silvi's model. The results agree and/or complement the state of the art.

1 Introduction

The valence theory of Lewis remains the basis for most modern ideas on the chemical bond. According to Lewis structures, there are bonding electron pairs in the valence shell of an atom in a molecule, and there are nonbonding electron pairs or lone pairs in the valence shell of many atoms in a molecule. From the topological analysis of the electron density, Bader *et al.* had extracted useful information about the bonding in a molecule. But, not much progress was made to reveal the location of these electron pairs [1].

According to Silvi *et al.* [2], the electron density alone does not easily reveal the consequences of the Pauli exclusion principle on the bonding. The work of several authors have produced a series of electron localization functions, which attempt to measure the Pauli repulsion by considering the Fermi hole. Hence, an alternative interpretation of these electron localization functions is to consider a system of fermions and a system of bosons with identical densities. The ground-state local kinetic energy of the non-interacting bosonic system is a lower bound to the local kinetic energy of the fermionic one. The excess local kinetic energy due to the Pauli principle is just the difference between the two. Where electrons are alone or form pairs of opposite spins, the Pauli principle has little influence on their behavior and they almost behave like bosons. In such regions the excess local kinetic energy has a low value. This identifies regions called attractors, every attractor consists of two electrons. There are three types: point, core and ring attractors. In this way, Silvi *et al.* is capable to locate and classify the electron pairs in organic molecules.

Nevertheless, in order to have this "non-interacting bosonic system", the magnetic momenta of the pairs of opposite spins are necessarily cancelling each other. Therefore, if the rotational energy barrier for the single bond in ethane has a magnetic origin, Silvi's model would not be able to explain it. The need to understand this molecular mechanism had driven chemists away from Silvi's most advanced model to semi-empirical ones. Currently, the origin of a rotational barrier in a C-C single bond has a wide range of explanations. The barrier is often attributed to: 1) torsional strains in the molecule, 2) steric strains, 3) charge transfer, exchange or electrostatic and 4) hyperconjugative interactions [3].

This is of a foremost interest because it has been found that the rotational speed of the bond reduces in the presence of an external magnetic field [4].

In the model used in this paper*, covalent bonds, lone pairs and core electrons will be detected by using the structures observed in Fig. 1, namely: the two separated spheres (*ts*), the torus (*t*) and the sphere in a sphere (*ss*) [5]. In the case of a single C-C bond, the magnetic momenta of the two bonding electrons are left to interact between each other. The C-C double bond would be two single bonds that consequently are locked for rotation. The C-C triple bond presents a lone pair (a torus) around its double bond structure and benzene presents interacting toroidal lone pairs, which are responsible for aromaticity.

Full count and location of the electron pairs forming different bonds, as well as, lone pairs is achieved. This was comparable or better than Silvi's model (the state of the art) [2]. The model/method presented in this paper: 1) confirmed Silvi's model electron count for certain molecules, 2) produced more information about missed electrons, not accounted by Silvi's model and shed light on the possible mechanism behind rotational barrier and aromaticity.

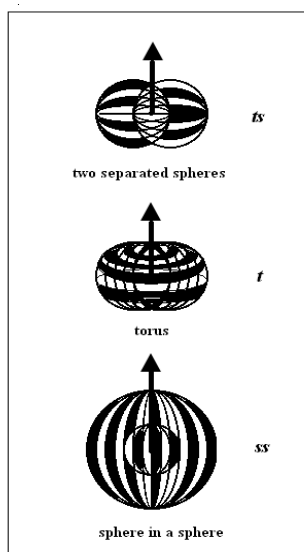*which has already been described in [5].

Fig. 1: Observables structures of the electron. The arrow represents its magnetic moment

2 Experimental

After observing the Laplacian of the electron density contour map of different hydrocarbon molecules, it was easy to identify C-C and C-H bonds and cut their silhouettes printed on paper. These silhouettes were weighted. The C-C or C-H bond lengths were used to calibrate the area measured in each bond. By this way, the bond area was calculated and it is reported in pm^2. An example of this process is in Fig. 4 for the C-H bond, and in Fig. 8 for the C-C bond. Then, these areas were correlated with their respective bond energies. A linear correlation was possible after dividing the bond area by a whole number, n. This whole number is interpreted as the number of electrons participating in the bond and it is reported on the right side of the molecule formula. These are observed in Figs. 2 and 3. This method has been sufficiently described in [5] and, in this paper, it was applied to the hydrocarbon molecules: ethane, ethene, ethyne and benzene. The contour map of the Laplacian of the charge density for C-H and C-C bonds in ethane, ethene and ethyne molecules are in [6]. Benzene C_6H_6 in [7] and C_2 is in [8].

2.1 Electron count

Fig. 2 shows that with n very close to 2, the C-H bond area linearizes against the bond energy in the molecules: ethane C_2H_6, 2; benzene C_6H_6, 2.01 and ethyne C_2H_2, 2. In the case of ethene C_2H_4 it is 1.824. Fig. 2 shows that n is exactly 2 in the case of C-C ethane and benzene, 8 in the case of dicarbon and 4 in the case of C-C ethyne. Ethene, however, presents 2.6 for the C-C bond in the plane of the molecule and 4 in the plane perpendicular to it and at the C-C axis.

The number of electrons involved in the C-H bond was very close to 2 regardless the class of C-H bond. The C-H
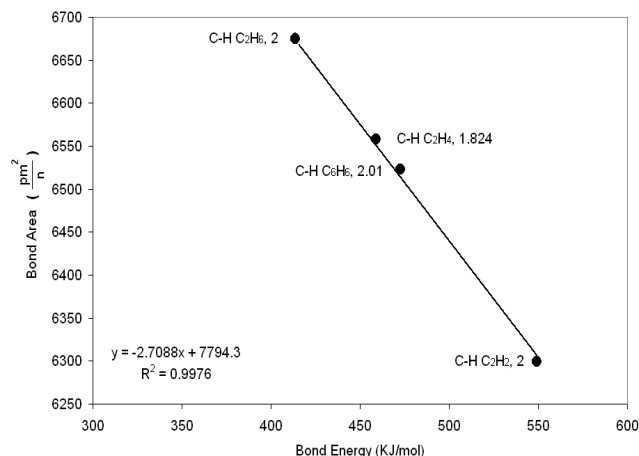

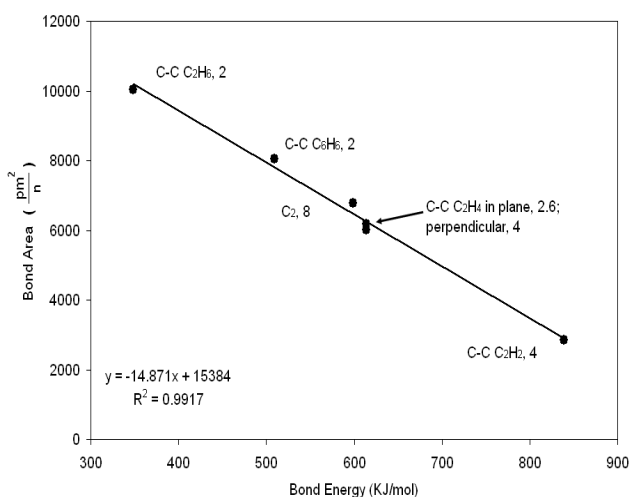Fig. 2: Bond area vs. bond energy for C-H bonds in different molecules.

Fig. 3: Bond area vs. bond energy for C-C bonds in different molecules.

bond that was far from this behavior was C-H ethene with 1.824. This deviation will be further discussed later. Thus, two electrons are involved in the C-H bond in the cases of ethane, ethyne and benzene.

Given that Fig. 3 provides the number of electrons involved in each C-C bond for these molecules, one is ready to do the full count of electrons in each molecule.

2.1.1 Ethane, C_2H_6

Figs. 2 and 3 inform that the C-H and C-C bonds have two electrons each. Hence, as it is observed in Fig. 4a, ethane has the expected electron count for each bond. This electron distribution coincides with the one presented by Silvi et al. (Fig. 4b) where the black circles are point attractors with two electrons each. Silvi's model put these attractors at the mid-point of the C-C bond and towards the hydrogen atom in the C-H bond. This is probably due to electronegativity dif-
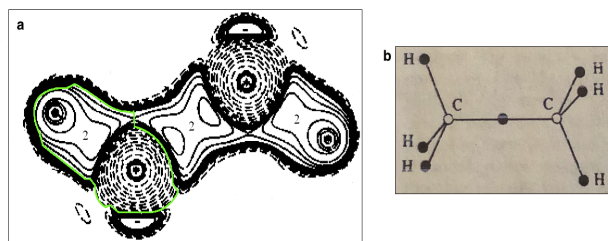
Fig. 4: **a)** Contour map of the Ethane molecule and its electron count. The green line shows how the C-H bond was cut. The C-C bond was also cut accordingly. Reprinted and adapted with permission from [6]. Copyright (1996) American Chemical Society. **b)** Silvi's et al. structure from [2]. It presents point attractors (black circles) alongside C-C and C-H bonds. It also shows core attractors (open circles) on the carbon atoms, used with permission of the publisher.

ferences between the bonding atoms. He also localized core attractors (open circles) on the carbon atoms. The model presented in this paper does not have that.

2.1.2　Rotational barrier

The ethane molecule presents one of the long standing problems in physical chemistry. This molecule has an energy barrier to its rotation. This barrier produces two types of conformers: the eclipsed and the staggered (see Fig. 5). The energy barrier between them is about 12 kJ/mol. Also, the C-C bond contracts from 153 pm in the staggered to 130 pm in the eclipsed conformer [3].

In between several explanations, the most favored ones are: 1) steric hindrance and 2) hyperconjugation. Although the steric effect is usually defined as the repulsion between C-H bonds or vicinal H atoms in the eclipsed conformation, the difference between torsional and steric strain is not clear. This is because they are not explicitly associated with a well-defined physical property.

Within the framework of natural bond orbital analysis, NBO, hyperconjugation is considered to be the source of the conformational preference of the molecule, by means of σ_{C-H} - $\sigma_{C-H}*$ vicinal interactions, rather than the electrostatic contribution or Pauli repulsion.

Most other explanations in the literature are given either in terms of orbital interactions or based on an energetic analysis of the problem. The discussion is far from over [3].

In the model presented in this paper, the electron is observed as the size of the whole bonding region. Given that the electron is also a tiny magnet, the interaction of the magnetic momenta between the two bonding electrons of the C-C bond is directly the cause of this torsional barrier and the differences in the C-C length between conformers.

Fig. 5 presents the two configuration and the magnetic

momenta of the two bonding electrons. In the eclipse conformer, these magnetic momenta are at an angle of 180 degrees (maximum magnetic attraction). This shortens the C-C bond to 130 pm. Upon rotation of one of the carbon atoms, the angle between electron's magnetic momenta decreases. At $180 - 60 = 120°$, a combination of distance between moments and the angle vanished this magnetic interaction. This lengthens the C-C bond (minimum magnetic attraction) in the staggered conformer.

The equation that describe the interaction between the two electron magnets is,

$$F = \frac{3\mu_0}{4\pi} \frac{m_e^2}{r^4} \cos\theta \qquad (1)$$

where μ_0 is the permeability of the free space, m_e is the electron magnetic moment, r is the distance between magnetic moments and θ is the angle between them. Mimicking the magnitude of the Ehrenfest forces acting on the C atoms, Fe(C) for diferent C-C distances presented in [3]. The change in magnetic force, equation (1), needed to explain the barrier at different C-C distances is presented in Fig. 6.

Given that there are no other energy barrier, it is believed that the bond rotation occurs in step between the carbon atoms in the bond. This means that once one carbon reached the weakening angle, the other rotates to reach 180° again. This mechanism would be consistent with a reduction in the rotation speed in the presence of an external magnetic field, which has been experimentally detected [4]. Silvi's model is simply incapable to reproduce this interaction because the bonding electrons' magnetic momenta are not free to interact in this way.[*]

2.2　Ethene, C_2H_4

Fig. 7a shows so far, the electron count extracted from the results in Figs. 2 and 3. Since $4(1.824) + 2.6 \approx 10$, a deficit of two electrons remains unexplained. However, the C-C electron count in the plane perpendicular to the molecular plane at the C-C axis gives exactly 4 (see Figs. 3 and 7b). This is, even though no indication of localization in this region is observed and these 4 electrons look to be in the same region of space (fused). This count probably means that the C-H electron count on the molecular plane, 1.824 is 2 in the plane perpendicular to it. Thus, a full electron count of this molecule is obtained. Coincidently, Silvi's model presents same electron count and localization. Two point attractors (4 electrons) at the plane perpendicular to the molecular plane for ethene: one over and the other under the molecular plane and point attractors (2 electrons each) for the four C-H bonds in the molecular plane (see Fig. 7c).

The C-C single bond results, already described for ethane, provide a way to understand the double bond. Simply, after

_____

[*]they are cancelling each other, completely coupled to obey the Pauli principle.
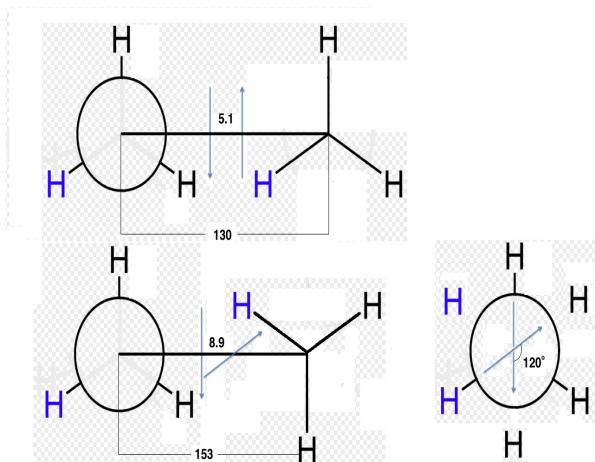
Fig. 5: Eclipsed (top) and Staggered (down) ethane conformers. The eclipsed conformer present the maximum magnetic attraction and shorter C-C bond length. Whereas, the staggered conformer has the lowest magnetic attraction and the longest C-C bond length. The extra projection shown down right is to present the angle between the two bonding electrons magnetic momenta in the staggered conformer. All distances are in pm.

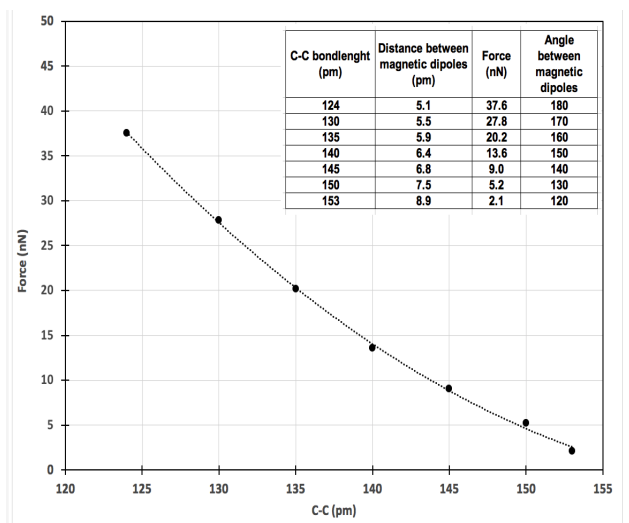| C-C bondlenght (pm)	Distance between magnetic dipoles (pm)	Force (nN)	Angle between magnetic dipoles
124	5.1	37.6	180
130	5.5	27.8	170
135	5.9	20.2	160
140	6.4	13.6	150
145	6.8	9.0	140
150	7.5	5.2	130
153	8.9	2.1	120

Fig. 6: Magnetic force between carbon atoms in the C-C bond for different C-C distances. The insert present the force values, distance between magnetic dipoles and angle assigned to each C-C distance. The integral of the curve is 12 kJ/mol.

the first single bond occurs, a second single bond in the C-C bond will lock any possibility for rotation. This is concurrent in both models presented here. Furthermore, Silvi's model does not present a point attractor in the line between the two carbons. Thus, the double bond looks like two out of line sigma bonds.
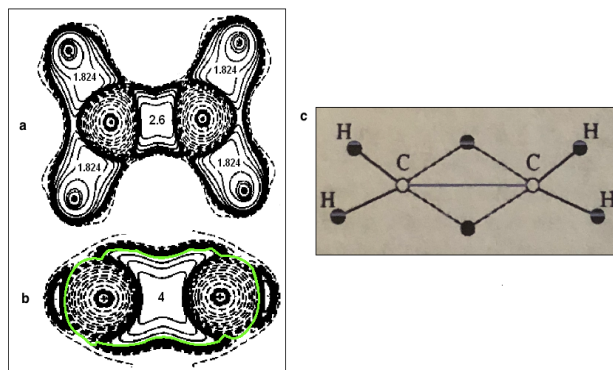

Fig. 7: Contour map of the Ethene molecule and its electron count. The C-C bond electron count at the plane of the nuclei (**a**) is different from the count at the perpendicular plane (**b**). The green line shows how the C-C bond was cut. Reprinted (adapted) with permission from [6]. Copyright (1996) American Chemical Society. **c**) Silvi's et al. structure from [2] used with permission of the publisher.
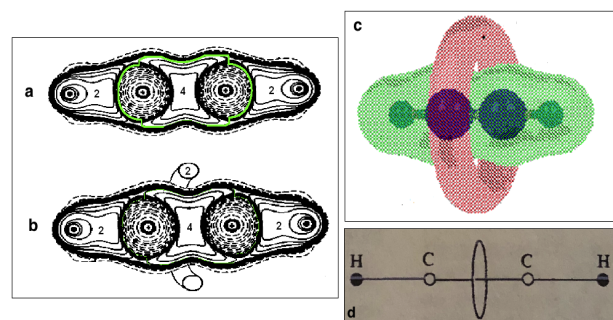

Fig. 8: **a**) Contour map of the ethyne molecule and its electron count, there is a lack of two electrons. The green line shows how the C-C bond was cut. **b**) These two electrons are fused in a toroidal lone pair around the C-C bond. Reprinted (adapted) from [6]. Copyright (1996) American Chemical Society. **c**) This structure has been observed in the molecular electrostatic potential of ethyne. This is from [9] used with permission of the publisher. **d**) Silvi's et al. structure from [2] used with permission of the publisher.

2.3 Ethyne, C_2H_2

Fig. 2 presents that C-H bond has two electrons in ethyne, Fig. 3 shows that the C-C bond has 4. Therefore, Fig. 8a presents a lack of two electrons. These two electrons will be bonded outside of the ethyne's C-C bond and at its midpoint, completely fused, producing a lone pair with a toroidal shape (see Fig. 8b). This has been observed in the molecular electrostatic potential of this molecule (see Fig. 8c [9]). This toroidal shape has also been noticed as a "ring attractor" in the electron localization function, $\eta(\mathbf{r})$, of this molecule in [6]. Concurrently, Silvi's structure also presents this ring attractor (2 electrons) and the point attractors for the C-H bonds, see Fig. 8d. But, it misses the other four electrons in the C-C bond. Fig. 8b depicts the complete electron count for the
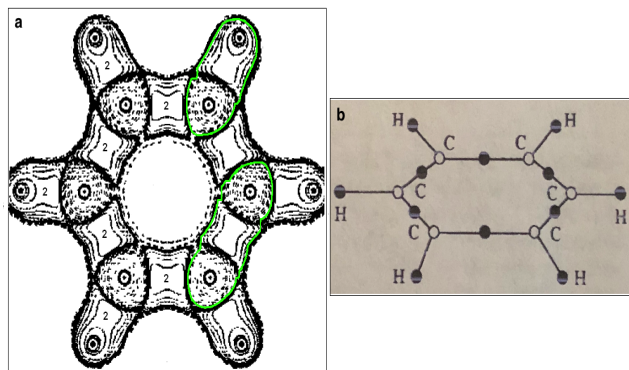
Fig. 9: **a**) Contour map of the Benzene molecule and its electron count. There is a lack of six electrons. These electrons are fused in two lone pairs at both sides of the C_6 ring. The green line shows how the bonds were cut. This is from [7] used under Creative Commons license. **b**) Silvi's et al. structure from [2] used with permission of the publisher.

ethyne molecule.

The evidence shows that the triple bond is a double bond with a lone pair. Also, the availability of two more bonding electrons would make this lone pair to disappear into a quadruple bond, which has been observed in dicarbon [5].

2.4 Benzene, C_6H_6

Figs. 2 and 3 show that the C-H and C-C bond in benzene have two electrons each. Fig. 9a presents the electron count for benzene. Silvi's structure (Fig. 9b) also depicts the same C-H and C-C electron count. None of these structures inform the whereabouts of the six remaining electrons. It is believed that they will go to two fused toroids (three electrons each) on both sides of the C_6 molecular plane. This is because that has been observed in the molecular electrostatic potential of benzene [9] (see Fig. 10). The aromatic stabilization energy for benzene is 120 kJ/mol [10], which is comparable to a weak chemical bond (for example F-F with 155 kJ/mol [5]). Thus, it is believed that these lone pairs act as such.

2.4.1 Aromaticity

In the customary view of aromaticity, an external magnetic field induces a molecular plane ring current in the delocalized π electrons of the aromatic ring. This current will produce its own magnetic field, which will go against the external magnetic field. This effect will deshield protons outside of the molecular plane. According to Fig. 10, there are three electrons in each toroidal lone pair; two of them are magnetically coupled and the third one will be uncoupled. The same structure occurs on the other side of the molecular plane. Therefore, they will magnetically attract across such plane (see Fig. 10). When an external magnetic field is imposed on the benzene molecule, these toroidal lone pair

structures will align their magnetic momenta against the external magnetic field naturally resisting to lose its original and more stable configuration. As in the customary explanation, this effect will deshield the protons outside of the molecular plane.
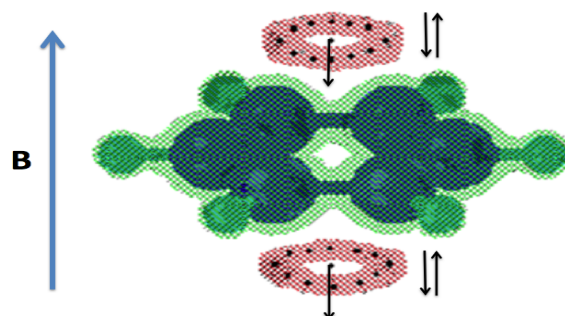

Fig. 10: Molecular electrostatic potential of benzene. The arrows depict the coupling of the three electrons in each lone pair. When an external magnetic field **B** is imposed, the magnetic moments of the two odd electrons aligned against it as shown. This is from [9] used with permission of the publisher.

3 Conclusions

A new experimental method to find the number of electrons shared in a chemical bond has been applied to selected hydrocarbon molecules. The information obtained is comparable and/or complements the state of the art. The total distribution of electrons in four fundamental hydrocarbons has been achieved. The long standing mystery of the ethane rotational barrier has been explained. The interaction between bonding electron magnets presents itself as fundamental to understand organic molecules.

Received on August 7, 2020

References

1. Bader R. F. W. Atoms in Molecules, a Quantum Theory. Clarendon Press, Oxford, 1990, p. 163.

2. Silvi B. and Savin A. *Nature*, 1994, v. 371, 683.

3. Cortés-Guzmán F., Cuevas G., Martín Pendás A. and Hernández-Trujillo J. *Phys. Chem. Chem. Phys.*, 2015, 17, 19021

4. Stavretis S. E. Probing Magnetic and Vibrational Properties of Molecular Compounds by Neutron Scattering, PhD dissertation, University of Tennessee, 2018. p. 135.

5. Yépez O. *Prog. Phys.*, 2019, v. 15, 3.

6. Bader R. F. W., Johnson S., Tang T. H. and Popelier P. L. A. *J. Phys. Chem.*, 1996, v. 100, 15398.

7. Bader R. F. W. and Keith T. A. *J. Phys. Chem.*, 1993, v. 99, 3683.

8. Chan W. and Hamilton I. P. *J. Phys. Chem.*, 1998, v. 108, 2473.

9. Gadre S. R., Bhadane P. K. *Resonance: Journal of Science Education*, 1999, v. 4, 14.

10. Suressh C. H. and Koga N. *J. Org. Chem.*, 2002, v. 67, 1965.

The Unpublished Feynman Diagram IIc

Oliver Consa

Department of Physics and Nuclear Engineering, Universitat Politècnica de Catalunya,
Campus Nord, C. Jordi Girona, 1–3, 08034 Barcelona, Spain.
E-mail: oliver.consa@gmail.com

Quantum Electrodynamics (QED) is considered the most accurate theory in the history of science. However, this precision is limited to a single experimental value: the anomalous magnetic moment of the electron (g-factor). The calculation of the electron g-factor was carried out in 1950 by Karplus and Kroll. Seven years later, Petermann detected and corrected a serious error in the calculation of a Feynman diagram; however, neither the original calculation nor the subsequent correction was ever published. Therefore, the entire prestige of QED depends on the calculation of a single Feynman diagram (IIc) that has never been published and cannot be independently verified.

1 Introduction

According to the Dirac equation, the value of the magnetic moment of the electron should be exactly one Bohr magneton. In 1947 it was discovered that the experimental value of the magnetic moment of the electron presented an anomaly of 0.1% with respect to the theoretical value [1] [2]. This anomaly was called the electron g-factor

$$\mu_e = g\mu_B = g\frac{e\hbar}{2m_e}\,. \tag{1}$$

Schwinger carried out the first theoretical calculation of the electron g-factor obtaining a value very similar to the experimental value. This value is known as the Schwinger factor [3]

$$g = 1 + \frac{\alpha}{2\pi} = 1.001162\,. \tag{2}$$

According to Quantum Electrodynamics (QED), the theoretical value of the electron g-factor is obtained by calculating the coefficients of a number series called the Dyson series [4]. When Feynman, Schwinger, and Tomonaga received the 1965 Nobel Prize for the development of QED, only the first two coefficients in the series had been calculated. The rest of the coefficients in the Dyson series were calculated many years later with the help of supercomputers

$$g = C_1\left(\frac{\alpha}{\pi}\right) + C_2\left(\frac{\alpha}{\pi}\right)^2 + C_3\left(\frac{\alpha}{\pi}\right)^3 + C_4\left(\frac{\alpha}{\pi}\right)^4 + C_5\left(\frac{\alpha}{\pi}\right)^5 \ldots \tag{3}$$

Each coefficient in the series requires the calculation of an increasing number of Feynman diagrams. The first coefficient in the Dyson series is the Schwinger factor and has an exact value of 0.5. The second coefficient was calculated in 1950 by Karplus and Kroll [6], who obtained a result of -2.973. This result was corrected seven years later by Petermann [8], who obtained a result of -0.328, almost 10 times lower than the previous calculation

$$g = 1 + \frac{1}{2}\left(\frac{\alpha}{\pi}\right) - 0,328\left(\frac{\alpha}{\pi}\right)^2 = 1,0011596\,. \tag{4}$$

The error was found in the calculation of the Feynman diagram IIc. According to the Karplus and Kroll original calculation, the value of diagram IIc was -3.178 while in the Petermann correction the value of diagram IIc was -0.564.
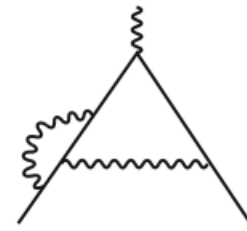

Fig. 1: Feynman diagram IIc.

The entire prestige of QED is based on its impressive level of precision of the electron g-factor. Currently QED allows the achievement of the electron g-factor with a precision of 12 decimal places of the theoretical value with respect to the experimental value

- 2008 Gabrielse's experimental value [13]:

  1.001 159 652 180 73(28) ;

- 2018 Kinoshita's theoretical value [14]:

  1.001 159 652 182 032(720) .

The calculation of the electron g-factor is based on the calculation of the second coefficient of the Dyson Series. The second coefficient of the Dyson series is based on the calculation of the Feynman diagram IIc. Therefore, the calculation of the Feynman diagram IIc performed by Karplus and Kroll in 1950 [6] can be considered the most important calculation in the history of modern physics.

Surprisingly, the original calculation of this diagram IIc turned out to be wrong and was corrected seven years after its publication. Inexplicably, both the original Feynman diagram IIc calculation and the subsequent correction have never been published, so the most important calculation in the history of modern physics cannot be independently verified.

2   Original calculation

2.1   Karplus and Kroll's paper

In 1949, Gardner and Purcell [5] published a new experimental result for the electron g-factor of 1.001146. In response, Karplus and Kroll performed the necessary calculations to obtain the second coefficient in the Dyson series.

In 1950, Karplus and Kroll [6] published a value of -2.973 for the second Dyson series coefficient and a new theoretical value of 1.001147 for the electron g-factor, in good agreement with the experimental data

$$g = 1 + \frac{\alpha}{2\pi} - 2.973\left(\frac{\alpha}{\pi}\right)^2 = 1.001147. \qquad (5)$$

The paper, published February 14 in the Physical Review Journal 77, consists of 14 pages full of complex mathematical calculations.

On the second page of the document, the authors indicate that to obtain the coefficient, it is necessary to calculate 18 Feynman diagrams grouped in five groups (I, II, III, IV and V). However, on pages 3 and 4, they argue that groups III, IV and V are not necessary. Therefore, it is only necessary to calculate seven Feynman diagrams, identified as I, IIa, IIb, IIc, IId, IIe, IIf. A lot of calculations are done between pages 4 and 11 that only serve to show that diagrams IIb and IIf are not necessary either. Therefore, it is only necessary to calculate five Feynman diagrams (I, IIa, IIc, IId, IIe).
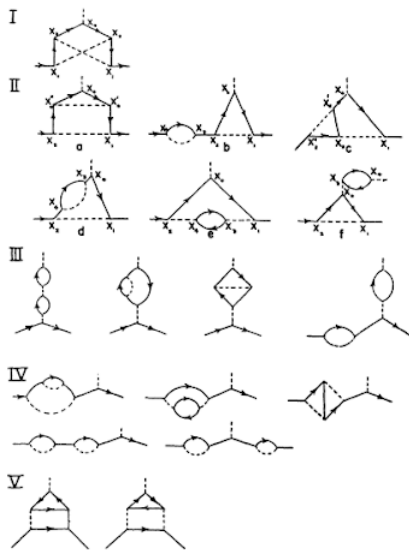

Fig. 2: Feynman diagrams.

The calculation of diagrams IIe (0.016) and IId (-0.090) are performed on pages 11 and 12 respectively. It follows that [6] "*The expressions for I, IIa and IIc become successively more complicated and very much more tedious to evaluate and cannot be given in detail here*". In other words, the complete calculation of three of the five diagrams was never

published. On page 13, the results of the three remaining diagrams are shown (I = -0.499, IIa = 0.778 and IIc = -3.178). Finally, page 14 of the paper presents the "Summary of Results" with the results of each of the five diagrams

$$C_2 = I + IIa + IIc + IId + IIe = -2,973. \qquad (6)$$

I	IIa	IIc	IId	IIe	Total
-0.499	0.778	-3.178	-0.090	0.016	-2.973

Table 1: Values of the five Feynman diagrams.

From the analysis of the results, it is evident that diagram IIc is the dominant diagram. Diagrams I and IIa are less relevant and practically cancel each other out. Diagrams IId and IIe are the only two diagrams whose calculations are included in the paper; however, their values are completely irrelevant.

The calculation of Feynman diagram IIc is made up of four components:

$$II_c = -\frac{323}{24} + \frac{31}{9}\pi^2 - \frac{49}{6}\pi^2 \ln(2) + \frac{107}{4}\zeta(3). \qquad (7)$$

Constant	π^2	$\pi^2 \ln 2$	$\zeta(3)$	Total
-13.458	33.995	-55.868	32.153	-3.178

Table 2: Value of the four components of Feynman diagram IIc.

The four components of IIc have abnormally high values (-13, 34, -55 and 32) which surprisingly compensate for each other, resulting in -3,178, an order of magnitude lower. It is not possible to say anything more about the calculation of diagram IIc because the complete calculation was never published.

The authors indicate that [6]: "*The details of two independent calculations which were performed so as to provide some check of the final result are available from the authors*". That is, the authors affirm that the calculations were carried out independently by two teams who obtained the same result, as a guarantee that the calculations were correct.

2.2   New experimental value

Six years after the publication of the Karplus and Kroll paper, Franken and Liebes [7] published new and more precise experimental data that showed a very different value for the electron g-factor (1.001165). This value was higher than the Schwinger factor, so the value of the second coefficient calculated by Karplus and Kroll not only did not improve the Schwinger factor, but made it worse. With the new experimental data, the value of the second coefficient in the series should have been +0.7 instead of -2.973.

Karplus and Kroll admitted that two independent calculations had not been carried out, so it was possible that there were errors in the calculations. According to Kroll [15]: "*Karplus and I carried out the first major application of that program, to calculate the fourth order magnetic moment, which calculation subsequently turned out to have some errors in it, which has been a perpetual source of embarrassment to me, but nevertheless the paper I believe was quite influential. (...) The errors were arithmetic (...) We had some internal checks but not nearly enough. (...) it was refereed and published and was a famous paper and now it's an infamous paper*".

The history of this correction is complex and confusing. We will now try to reconstruct this story from the published papers and quotes from its protagonists.

3 The history of the correction

3.1 Petermann's numerical calculation

Petermann was the first person to identify an error in the original calculation of Karplus and Kroll. He performed a numerical analysis of the five Feynman diagrams and he found that the solution of diagram IIc was clearly wrong, since its value was outside the limits. The rest of the diagrams were within limits [9]: "*The numerical results for the terms I, IIa, IIc, IId, IIe in the work by Karplus and Kroll have been checked by rigorous upper and lower bounds. Whereas every other term fell well between these bounds, agreement could not be obtained for diagram IIc. (...) The numerical value for this term has been found to satisfy IIc = -1.02 +/- 0.53*".

Petermann published a second paper where he adjusted his calculations [10]: "*the diagram IIc is found to satisfy IIc = -0.60 +/- 0.11 in contradiction with the value -3.18 given by the previous authors*".

Between the publication of these two papers, Petermann communicated privately to Sommerfield the result of another calculation [11]: "*Note added in proof. Petermann has placed upper and lower bounds on the separate terms of Karplus and Kroll. He finds that their value for IIc does not lie within the appropriate bounds. Assuming the other terms to be correct, he concludes that the result is -0.53 +/- 0.37*".

Petermann worked for three months following a numerical methodology that allowed him to narrow the margin of error in diagram IIc. Surprisingly, fourteen days after his third numerical calculation, he made an unexpected change in his methodology and published the exact analytical calculation, with no margins of error.

The articles published by Petermann on the calculation of the Feynman diagram IIc are summarized in Table 3.

3.2 Sommerfield and the Green's functions

After the publication of the new experimental value by Franken and Liebes [7], Schwinger commissioned a 22-year-old student named Sommerfield to redo the Kroll and Karplus

Date	IIc	Method	Publication
28/5	-1.02 +/- 0.53	Numerical	Nuclear Phys. 3
1/7	- 0.53 +/- 0.37	Numerical	Phys. Rev. 107, Note added in proof. Private comm. with Sommerfield
3/8	-0.60 +/- 0.11	Numerical	Nuclear Phys. 5
17/8	-0.564	Analytical	Helvetica Physica Acta 30

Table 3: Petermann's publications.

calculations. Schwinger proposed using his own method based on Green's functions instead of using Feynman diagrams.

According to Sommerfield's testimony [16]: "*Julian assigned us three problems, one of which involved the anomalous magnetic moment (...). At my meeting with him, he suggested that I continue the calculation of the anomalous magnetic moment to the next fourth order (...). Schwinger wanted me to use the other method, while respecting gauge invariance at every step. Many years later Roy Glauber told me that the faculty was not entirely happy that a graduate student had been given such a problem*".

In May 1957, Sommerfield sent a two-page paper to the Physical Review Journal where he published his results [12]: "*The fourth-order contribution to the moment is found to be* −0.328 *(..) Thus the result is 1.0011596*". This new theoretical value of the electron *g*-factor was in good agreement with the new experimental value of Franken and Liebes.

As Schwinger states [18]: "*Interestingly enough, although Feynman-Dyson methods were applied early [by Karplus and Kroll], the first correct higher order calculation was done by Sommerfield using [my] methods*".

The second coefficient of the Dyson series calculated by Sommerfield consisted of four components, the same as the original result for Karplus and Kroll, but with very different values:

[K&K]

$$C_2 = -\frac{2687}{288} + \frac{125}{36}\pi^2 - 9\pi^2 \ln(2) + 28\zeta(3) = -2.973. \quad (8)$$

[Sommerfield]

$$C_2 = \frac{197}{144} + \frac{1}{12}\pi^2 - \frac{1}{2}\pi^2 \ln(2) + \frac{3}{4}\zeta(3) = -0.328. \quad (9)$$

Sommerfield's paper does not include the calculations performed, but the author states that [11]: ""*The present calculation has been checked several times and all of the auxiliary integrals have been done in at least two different ways*". As a guarantee that the calculations were correct.

	Const.	π^2	$\pi^2 \ln(2)$	$\zeta(3)$	Total
K&K	-9.329	34.269	-61.569	33.656	-2.973
Pet.	1.368	0.822	-3.421	0.901	-0.328
Diff.	10.697	-33.447	58.148	-32.754	2.645

Table 4: Comparative components of C_2.

In 1958, Sommerfield published his g-factor calculations in the Annals of Physics [12] as part of his doctoral thesis. If we analyze his extensive 32-page paper, we verify that he used Green's functions instead of Feynman diagrams. For this reason, the calculation of the enigmatic Feynman diagram IIc does not appear in this paper.

In the third volume of "Particles, Sources, and Fields" published in 1989 [3], Schwinger devoted more than 60 pages to a detailed calculation of the second coefficient of Dyson series getting exactly the same result, but, once again, using Green's functions instead of Feynman diagrams.

In his 1957 paper, Sommerfield also states that [11]: "*The discrepancy has been traced to the term I and IIc of Karplus and Kroll*". This statement about the origin of the error cannot be deduced from Sommerfield's calculations, since he used Green's functions instead of Feynman diagrams. So Sommerfield had to receive this information from other sources (Petermann, Karplus or Kroll).

3.3 Petermann's definitive correction

The definitive solution to the problem was presented in 1957 by Petermann in a paper published in the Swiss journal Helvetica Physica Acta [8]. Although the paper was signed by a single author, actually the result was obtained by consensus between the results of the Petermann's numerical analysis, the Sommerfield calculation of C_2 using Green's functions and the correction of the Feynman diagrams carried out by Kroll himself. Petermann acknowledges that the result was obtained by consensus [8]: "*The new fourth order correction given here is in agreement with: (a) The upper and lower bounds given by the author. (b) A calculation using a different method, performed by C. Sommerfield. (c) A recalculation done by N. M. Kroll and collaborators*".

The article was signed by a single author due to an internal conflict between the researchers. As Sommerfied recalls [16]: "*In the meantime Schwingerian Paul Martin had gone to the Niels Bohr Institute in Copenhagen and had spoken to Andre Petermann, a postdoc with the Swedish theoretician Gunnar Kallen. Martin told Petermann about my work (...) In the end, however, after both of our calculations were completely finished they were in agreement with each other but not with Karplus and Kroll. We agreed to cite each other's work when published. However, Schwinger and Kallen had had a somewhat acrimonious discussion (...) and Kallen had*

forbidden Petermann to mention my work. Petermann's apology to me was profuse".

The Petermann final result for the electron g-factor was identical to the Sommerfield result published three months earlier

$$C_2 = \frac{197}{144} + \frac{1}{12}\pi^2 - \frac{1}{2}\pi^2 ln(2) + \frac{3}{4}\zeta(3) = -0.328. \quad (10)$$

In the paper, Petermann states that: "*We have performed an analytic evaluation of the five independent diagrams contributing to this moment in fourth order. The results are the following (I = -0.467, IIa = 0.778, IIc = -0.564, IId = -0.090, IIe = 0.016, Total = -0.328). Compared with the values in their original paper by Karplus and Kroll, one can see that two terms were in error: I differs by 0.031 and IIc differs by 2.614*".

I	IIa	IIc	IId	IIe	Total
-0.467	0.778	-0.564	-0.090	0.016	-0.328

Table 5: Corrected values of the five Feynman diagrams.

Comparing the results of the calculations of the Feynman IIc diagram carried out by Karplus and Kroll with the Petermann calculations we observe the following:

[K&K]

$$II_c = -\frac{323}{24} + \frac{31}{9}\pi^2 - \frac{49}{6}\pi^2 ln(2) + \frac{107}{4}\zeta(3) \quad (11)$$

[Petermann]

$$II_c = -\frac{67}{24} + \frac{1}{18}\pi^2 + \frac{1}{3}\pi^2 ln(2) - \frac{1}{2}\zeta(3) \quad (12)$$

The calculation of each of the four factors in diagram IIc is shown in the following table:

	Const.	π^2	$\pi^2 ln(2)$	$\zeta(3)$	Total
K&K	-13.458	33.995	-55.868	32.153	-3.178
Pet.	-2.791	0.548	2.280	-0.601	-0.564
Diff.	10.667	-33.447	58.148	-32.754	2.614

Table 6: Comparative components of Feynman diagram IIc.

The corrections are huge, one or two orders of magnitude for each component of diagram IIc. We cannot know the origin of these discrepancies because the correction calculations were also not published.

4 Summary

The calculation of the Feynman diagram IIc can be considered the most important calculation in the history of modern physics. However, the history of this calculation is surrounded by errors and inexplicable coincidences.

- The original calculation of the Feynman diagram IIc published in 1950 was wrong.

- Karplus and Kroll stated that the calculation had been performed by two teams independently. This statement was made to give guarantees about the validity of the calculations, and yet it turned out to not be the case.

- Despite having published a wrong result, the prestige of Karplus and Kroll was not affected at all. On the contrary, both enjoyed brilliant careers full of awards and recognition for their professional achievements.

- The Karplus and Kroll miscalculation was consistent with the experimental value previously published by Gardner and Purcell, even though that experimental value was also wrong.

- The error in the calculation was not reported until seven years after its publication.

- The error in the calculation was detected just when a new experimental value was published by Franken and Liebes. The corrected theoretical value also coincided with the new experimental value.

- Neither the original calculation of the Feynman diagram IIc nor its subsequent correction has been published to date.

Received on September 7, 2020

References

1. Nafe J. E., Nelson E. B., Rabi I. I. The Hyperfine Structure of Atomic Hydrogen and Deuterium. *Phys. Rev.*, 1947, v. 71 (12), 914–915.

2. Breit G. Does the Electron Have an Intrinsic Magnetic Moment? *Phys. Rev.*, 1947, v. 72 (10), 984–984.

3. Schwinger J. On Quantum-Electrodynamics and the Magnetic Moment of the Electron. *Phys. Rev.*, 1948, v. 73 (4), 416–417.

4. Dyson F. The Radiation Theories of Tomonaga, Schwinger, and Feynman. *Phys. Rev.*, 1949, v. 75 (3), 486–502.

5. Gardner J. H., Purcell E. M. A Precise Determination of the Proton Magnetic Moment in Bohr Magnetons. *Phys. Rev.*, 1949, v. 76 (8), 1262–1263.

6. Karplus R., Kroll N. Fourth-order corrections in quantum electrodynamics and the magnetic moment of the electron. *Phys. Rev.*, 1950, v. 77 (4), 536–549.

7. Franken P., Liebes S. Magnetic Moment of the Proton in Bohr Magnetons. *Phys. Rev.*, 1956, v. 104 (4), 1197–1198.

8. Petermann A. Fourth order magnetic moment of the electron. *Helvetica Physica Acta*, 1957, v. 30, 407–408.

9. Petermann A. Magnetic moment of the electron. *Nucl. Phys.*, 1957, v. 3, 689–690

10. Petermann A. Magnetic moment of the electron. *Nucl. Phys.*, 1958, v. 5, 677–683

11. Sommerfield C. M. Magnetic Dipole Moment of the Electron. *Phys. Rev.*, 1957, v. 107 (1), 328–329.

12. Sommerfield C. M. The Magnetic Moment of the Electron. *Annals of Physics*, 1958, v. 5, 26–57.

13. Hanneke D., Fogwell S., Gabrielse G. New Measurement of the Electron Magnetic Moment and the Fine Structure Constant. *Phys. Rev. Lett.*, 2008, v. 100 (12), 120801–120805.

14. Aoyama T., Kinoshita T., Nio M. Revised and improved value of the QED tenth-order electron anomalous magnetic moment. *Phys. Rev. D.*, 2018, v. 97 (3), 036001.

15. Kroll. N. Interview with Finn Aaserud: Interview conducted on 28 June 1986. Niels Bohr Library & Archives, American Institute of Physics, College Park, MD, USA, 1986. https://www.aip.org/history-programs/niels-bohr-library/oral-histories/28394

16. Sommerfield C. M. Schwingerians. Julian Schwinger Centennial Conference, 2019, pp. 207–211. Transcript of the video lecture recorded.

17. Schwinger J. Particles, Sources, and Fields. Vol. III. Addison-Wesley, Advanced Book Classics, 1989.

18. Mehra J., Milton K. A. Climbing the Mountain: The Scientific Biography of Julian Schwinger. Oxford University Press, 2000, p. 117.

————————

A Model of the Universe Expanding at a Constant Speed

Rostislav Szeruda

Roznov p.R. 75661, Czech Republic. E-mail: rostislav.szeruda@seznam.cz

This article deals with the possibility of finding an alternative model to the expanding universe model which can be in accordance with our astronomical observations. This is considered an easy but not usual model of closed universe with $k = 1$, $\Lambda = 0$ and $q = 0$ which provides that mass of this universe is not constant but stepwise increasing.

1 Basic ideas and existing work

This article is based on four basic ideas:

1. Model of the universe expanding at a constant speed [3]. Such a model of the universe is not by itself consistent with observation. We observe that the rate of expansion of our universe is accelerating.

2. The idea that the universe may be a black hole is dealt with in [2].

3. The universe was born from a single quantum of energy. The mass of the universe, its size, and the instant speed of particles with a non-zero rest mass are interrelated. The idea was inspired by the book [1].

4. The relative particle size shrinking. This effect makes it seem to us that the universe is not expanding at a constant speed, but that the speed of its expansion is increasing. This idea is new.

2 Constant speed expanding universe

We assume that there is no difference between what our universe is and how it appears to us. But is that true? Let us imagine that we are in a room the walls of which are expanding, and we are shrinking just as quickly at the same time. How would this room seem to us?

Let us consider a universe expanding at a constant rate. The elementary particles try to move at the maximum possible speed. The speed of expansion of the universe is a limitation of the instantaneous velocity of the elementary particles within. Thus, particles with zero rest mass (photons) can move as fast as the universe expands:

$$\dot{a} \equiv c \tag{1}$$

where:

$\dot{a}$ – speed of the universe expansion

c – speed of light in vacuum.

Further, let's suppose that particles with non zero rest mass have a tendency to move at the speed of light in vacuum too but due to their non zero rest mass they are not able to achieve that speed. The more their speed gets closer to the speed of light in vacuum, the higher their mass becomes and prevents them from moving faster.

Let's have a model of the universe described by Friedmann equations:

$$3\left(\frac{\dot{a}}{a}\right)^2 + \frac{3kc^2}{a^2} - \Lambda c^2 = 8\pi G\rho \tag{2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3c^2}\left(\rho c^2 + 3p\right) + \frac{\Lambda c^2}{3} \tag{3}$$

where:

G – gravitational constant

$\rho = \rho(t)$ – matter density in universe

$p = p(t)$ – pressure in the universe

a – expansion factor of the universe

$\ddot{a}$ – acceleration of the universe expansion ($\ddot{a} = 0$ for the model)

k – parameter of the universe curvature

Λ – Einstein cosmological constant.

Let's consider a Riemann space with a positive curvature, where:

1. $k = 1$.
2. $\Lambda = 0$.

The Friedmann equations are simplified to:

$$\rho = \frac{3c^2}{4\pi Ga^2} \tag{4}$$

$$p = -\frac{1}{3}c^2\rho. \tag{5}$$

The density of the universe is then inversely proportional to the square of the expansion factor a. It means that the linearly expanding universe is possible only on the condition that its mass is not constant but it rises proportionally to a. The more matter the universe contains, the larger it becomes and vice versa.

For a closed universe ($k = 1$), we can call the expansion factor a as the radius of the universe. Its volume is an elementary inter-sphere with surfaces $4\pi a^2 \sin^2 \psi$ and its thickness $a\,d\psi$ ($0 \le \psi \le \pi$). We get it by integration:

$$V = a^3\, 4\pi \int_0^\pi \sin^2 \psi \, d\psi = 2\pi^2 a^3. \tag{6}$$

The universe mass is then given by the equation:

$$M_v = \frac{3\pi c^2 a}{2G}. \tag{7}$$

3 Initial parameters of the universe

Consider that the universe didn't begin its existence with all of the matter contained therein today, but was born from a single energy quantum M_0 in a space of the size of the minimal quantum packet:

$$a_0 = \frac{\hbar}{2M_0 c} = \frac{G\hbar}{3\pi a_0 c^3}. \tag{8}$$

Thence

$$a_0 = \sqrt{\frac{G\hbar}{3\pi c^3}} \cong 5.26 \times 10^{-36}\,\text{m}. \tag{9}$$

The minimum time interval then:

$$t_0 = \frac{a_0}{c} = \sqrt{\frac{\hbar G}{3\pi c^5}} \cong 1.76 \times 10^{-44}\,\text{s}. \tag{10}$$

The first quantum mass M_0 of the universe is given by the relation:

$$M_0 = \sqrt{\frac{3\pi \hbar c}{4G}} \cong 3.34 \times 10^{-8}\,\text{kg}. \tag{11}$$

The universe we describe here resembles a black hole. Its size is directly proportional to the amount of matter it contains:

$$a_\bullet = \frac{2GM_\bullet}{c^2} \tag{12}$$

$a_\bullet$ – radius of a black hole (Schwarzschild radius, horizon of events)

$M_\bullet$ – mass of a black hole.

The mass of the first quantum of the universe M_0 is big enough to create a black hole, because the minimum mass of a black hole is given by Planck's relationship:

$$M_{\bullet_0} = \sqrt{\frac{\hbar c}{4G}} \cong 1.09 \times 10^{-8}\,\text{kg}. \tag{13}$$

Thus, the initial quantum was below the event horizon, which began at a distance given by the minimum size of the black hole:

$$a_{\bullet_0} = \sqrt{\frac{G\hbar}{c^3}} \cong 1.62 \times 10^{-35}\,\text{m}. \tag{14}$$

A black hole of this mass is characterized by temperature:

$$T_{\bullet_0} = \frac{\hbar c^3}{8\pi k G M_{\bullet_0}} = 1.13 \times 10^{31}\,\text{K}. \tag{15}$$

4 The evolution of the universe

Let the mass of the universe be varied by quanta corresponding to the mass of the first quantum M_0. Then the size of the universe will change in discrete values, and the passage of time won't be continuous, but it will flow in elementary jumps:

$$M_v = nM_0 \tag{16}$$

$$a = na_0 \tag{17}$$

$$t = nt_0 = n\frac{a_0}{c} \tag{18}$$

where:

n – natural number higher then zero.

The space where the initial quantum can occur is limited by the expansion function of the universe a. As the mass of the universe starts to grow, a will increase and matter will have more space to be located and to move. The total energy of the universe is permanently zero.

The universe can have zero total energy if the total gravitational potential energy that holds all its components together is negative and in absolute value is exactly equal to the sum of all positive energy in the universe contained in the masses and movements of the particles.

The matter growth within the universe does not occur by locally emerging new matter, but by increasing the velocity of motion of the initial quantum of energy to a speed close to the speed of light in vacuum:

$$M_v = \alpha_{v_m} M_0 = \frac{M_0}{\sqrt{1 - \frac{v_m^2}{c^2}}} = nM_0 \tag{19}$$

where:

v_m – instant speed of all elementary particles with non zero rest mass. Consider that this speed is the same for all the quantum of energy in the universe. However, the resulting velocity of the particles made up of these quanta appear to be slower as the quantum of energy can move back and forth through space.

The instant speed of particles with rest mass is given by:

$$v_m = \dot{a}\sqrt{1 - \frac{1}{n^2}} = c\sqrt{1 - \frac{1}{n^2}}. \tag{20}$$

The older the universe is, the closer the instant speed of particles with a non zero rest mass is to the universe expansion speed.

$$\lim_{n \to \infty} v_m = c.$$

At the present time, the two values are not practically distinguishable.

Considering quanta of energy as moving one-dimensional objects, their size should appear smaller due to relativistic contraction:

$$l = \frac{l_0}{\alpha_{v_m}} = l_0\sqrt{1 - \frac{v_m^2}{c^2}}. \tag{21}$$

The inner observer doesn't know that he is shrinking together with his entire planet, his solar system or his galaxy, at the same time that the universe itself expands. Because he measures the expansion of the universe in comparision with himself, it will seem to him that the universe is expanding faster than it actually is. Due to the contraction of distance, the gravitational force will appear to him stronger. He will attribute it to the greater mass of interacting objects.

Therefore, from the perspective of the internal observer, the size and mass of the universe will appear:

$$a_i = \alpha_{v_m}^2 a_0 = n^2 a_0 \tag{22}$$

$$M_i = \alpha_{v_m}^2 M_0 = n^2 M_0 \,. \tag{23}$$

The fact, that the universe expands with speed $\dot{a} = c$ perpendicular to our three-dimensional space and to all speed vectors in it, can be expressed by adding an imaginary mark before the value of the expansion speed. Generally, we can express the speed of a mass object w this way:

$$w = v + \iota\dot{a} \tag{24}$$

where:

v – an object speed in our three-dimensional space.

The square of w can be expressed in the form:

$$w^2 = v^2 - \dot{a}^2 \left(1 - \frac{v^2}{c^2}\right) . \tag{25}$$

Now the Einstein relativistic coefficient α gets the more general form:

$$\begin{aligned}
\alpha_w &= \frac{1}{\sqrt{1 - \frac{w^2}{c^2}}} = \alpha_v \alpha_{\dot{a}} \\
&= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{1}{\sqrt{1 + \frac{\dot{a}^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{1}{\sqrt{2}} \,.
\end{aligned} \tag{26}$$

The first coefficient α_v in the relation (26) is the standard form of Einstein coefficient α. The second coefficient $\alpha_{\dot{a}}$ is related with the speed of the universe expansion and it is constant. So the universe will appear to us $\sqrt{2}$ times bigger but not more massive.

$$a_i = \sqrt{2}n^2 a_0 \tag{27}$$

$$M_i = \frac{\sqrt{2}n^2 3\pi c^2 a_0}{2\sqrt{2}G} = n^2 M_0 \,. \tag{28}$$

The universe density will seem to be equal with the critical density:

$$\rho_i = \frac{3c^2}{4\pi G \rho (\frac{a}{\alpha_{\dot{a}}})^2} = \frac{3\dot{a}^2}{8\pi G a^2} = \frac{3H^2}{8\pi G} = \rho_k \tag{29}$$

where:

H – Hubble constant:

$$H \equiv \frac{\dot{a}}{a} = \frac{c}{a} \,. \tag{30}$$

The density of the universe, in case of inner observer, thus seems to be equal to the critical density. It corresponds to our observation. In contrast to the inflation model it happens not only effectively. Therefore, the entire universe appears to be non-curved - flat, even though it is closed.

5 The universe pressure

The change of the internal energy of the universe corresponds with the change of its energy. The universe can't exchange heat with its surroundings. Then the first theorem of thermodynamics has an easy form by which we can express a change of the universe energy as:

$$dU = -p\,dV = c^2\,dM \,. \tag{31}$$

Mass movement in the direction of the expansion of the universe and its rise with time induce a force, which has size:

$$F = \iota^2 \frac{dM}{dt} c = -\frac{3\pi c^4}{2G} \,. \tag{32}$$

This force acts on the surface:

$$S = 6\pi^2 a^2 \,. \tag{33}$$

This creates a pressure that is already known from the relation (5):

$$p = \frac{-c^4}{4G\pi a^2} = -\frac{1}{3} c^2 \rho \,. \tag{34}$$

The pressure in the universe is negative at a positive energy density. However, matter and radiation create positive pressure. It thus appears rather a local phenomenon operating in three-dimensional space, which has no effect on the four-dimensional universe as a whole.

6 The universe age and mass

Three-dimensional black holes radiate energy from their horizon into the surrounding space. The horizon of a black hole bound up to the universe produces radiation which is moving on the surface of a four-dimensional sphere and remains part of it. As the universe expands, it cools down in such a way that its temperature corresponds to the current temperature of the black hole horizon.

The temperature of the radiation emitted at the beginning of the universe is now the same as the temperature of the radiation from the event horizon. The universe thus appears as the interior of the black body, where the radiation density is given by:

$$U = \frac{\pi^2}{15} \frac{(kT)^4}{(\hbar c)^3} \,. \tag{35}$$

For the temperature of the relict radiation $T_r = 2.726\,\text{K}$ results the energy density $U \cong 4.18 \times 10^{-14}\,\text{J m}^3 \cong 0.26\,\text{eV cm}^3$

out of the relation (34). This corresponds to the measured value of the density of the relict radiation $0.25\,\text{eV}\,\text{cm}^3$.

If the temperature of the universe at its beginning corresponded to the temperature of the black hole horizon according to the relation (15), today it should correspond to the temperature:

$$T_{\bullet_n} = \frac{\hbar c^3}{8\pi k G M_{\bullet_n}} = \frac{T_{\bullet_0}}{n}. \tag{36}$$

If the temperature of the relict radiation $T_r = 2.726\,\text{K}$ corresponds to the present temperature of the universe and at the same time to the temperature of the radiation from the early universe:

$$n \cong 4.14 \times 10^{30}. \tag{37}$$

The size of the universe (from the perspective of an inner observer) is then:

$$a_i = \sqrt{2}n^2 a_0 \cong 1.27 \times 10^{26}\,\text{m}. \tag{38}$$

The Hubble constant H is then according to (29):

$$H = \frac{c}{a} \cong 2.35 \times 10^{-18}\,\text{s}^{-1} \cong 72.63\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}. \tag{39}$$

This value is consistent with the value of the Hubble constant determined in 2018: $H = 73.52 \pm 1.62\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$. The actual age of the universe is therefore:

$$t = \frac{1}{H} \cong 13.5 \times 10^9\,\text{yrs}. \tag{40}$$

The mass of the universe (from the point of view of an internal observer) is:

$$M_i = n^2 M_0 \cong 5.72 \times 10^{53}\,\text{kg}. \tag{41}$$

7 Visible and invisible matter

We already know how the mass and size of the universe as a whole changes. How can the mass and size of its parts change? The mass of all objects has to change, for an internal observer, according to:

$$m_2 = m_1 \frac{t_2}{t_1} = m_1 \frac{n_2^2}{n_1^2} \tag{42}$$

where:

m_1 – object mass at time t_1 ($\sim n_1^2$)

m_2 – object mass at time t_2 ($\sim n_2^2$).

The fact that this relationship is true can be seen in the motion of matter around the centers of galaxies. Outside the galaxy, the mass should move with velocity according to the standard model (see curve A in Fig. 1)

$$v^2 = \frac{GM_g}{r} \tag{43}$$

where:

M_g – galaxy mass

r – distance from the galaxy center.

If the mass at the edge of the galaxies is pulled away from the center of the galaxy due to the universe expansion and grows with distance ($M_g(r) \sim r$) according to the relation (42), although most of this mass cannot be observed, their velocity around the galaxy's gravitational center remains the same in Fig .1 – the rotation curve becomes flat from a certain distance from the center.
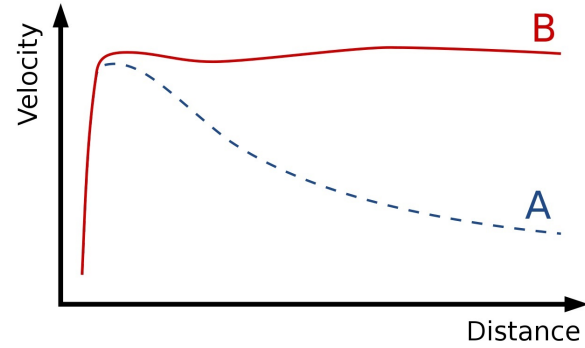

Fig. 1: Dependence of orbital velocity on distance from center of galaxy

The relation (42) describes the total amount of matter (perceivable or non-perceivable) that increases depending on space-time expansion. So what about the perceivable matter? If the relation (42) also applies to photons, and we still observe a redshift, this means that the first quantum of energy must be fragmented into a larger number of smaller quanta. For photons:

$$m_{f2} = m_{f1} \frac{n_{f1}}{n_{f2}} \frac{t_2}{t_1} \tag{44}$$

where:

n_{f1} – number of photons at time t_1

n_{f2} – number of photons at time t_2

m_{f1} – photon mass at time t_1

m_{f2} – photon mass at time t_2.

If the universe with temperature T_1 at time t_1 contained n_{f1} particles then it will have at time t_2 temperature T_2 and will contain n_{f2} particles:

$$\frac{n_{f2}}{n_{f1}} = \frac{p_2 V_2}{p_1 V_1} \frac{T_1}{T_2} = \frac{a_2}{a_1} \frac{T_1}{T_2} = \frac{t_2}{t_1} \frac{T_1}{T_2} = \frac{n_2^3}{n_1^3}. \tag{45}$$

After insertion into (44):

$$m_{f2} = m_{f1} \frac{T_2}{T_1} = m_{f1} \frac{n_1}{n_2} \tag{46}$$

$$\lambda_2 = \lambda_1 \frac{T_1}{T_2} = \lambda_1 \frac{n_2}{n_1}. \tag{47}$$

As the temperature of the universe is decreasing, the mass of the photons has to decrease too. The radiation on the way through the universe gets colder, but the number of observable photons increases – as if the universe in the past contained the same amount of matter as today.

The particulate mass with non-zero rest mass will grow by (42) but simultaneously their wavelength will lengthen according to (47). Their mass is then given:

$$m_2 = m_1 \frac{T_2}{T_1} \frac{t_2}{t_1} = m_1 \frac{n_2}{n_1} \tag{48}$$

where:

m_1 – mass of a "cold" particle at time t_1

m_2 – mass of a "cold" particle at time t_2.

As the mass of the universe increases, the number of quanta energies increases faster and thus their energy decreases. The smallest quanta of energy now have mass:

$$M_{0n} = \frac{n^2 M_0}{n^3} = \frac{M_0}{n} \cong 8.08 \times 10^{-39} \, \text{kg} \,. \tag{49}$$

Relationships (46) and (48) describe observable mass. This is obviously lesser than the mass objects should have by the equation (42). Mass corresponding to the difference we can't directly observe, but we can observe its gravitational effect. The matter we name: "dark matter". This is the "missing" matter around the galaxies.

8　Observable quantity of energy

A standardized wave packet is related with the whole universe and it moves in direction of the universe expansion [4]:

$$|\psi(a;t)|^2 = \frac{1}{\sqrt{2\pi} \, \Delta a_t} \exp\left[-\frac{(a - ct)^2}{2(\Delta a_t)^2} \right]. \tag{50}$$

The wave packet related to the universe shows a dispersion which causes it to seem higher. For as much that the mass of the universe increases linearly with time, the dispersion is independent of time:

$$\Delta a_t = \sqrt{(\Delta a_0)^2 + \left(\frac{\Delta(m_0 \dot{a})}{m} t \right)^2}$$

$$= \sqrt{a_0^2 + \left(\frac{m_0 c t_0}{m_0} \right)^2} = a_0 \sqrt{2} \,. \tag{51}$$

This result is in agreement with $\alpha_{\dot{a}} = 1/\sqrt{2}$ from the relation (26). The amplitude of this wave package relative to a_0 is then:

$$|\psi(a = ct; t)|^2 \, a_0 = \frac{1}{2\sqrt{\pi}} \cong 0.282 \,. \tag{52}$$

It means that if the universe size is a, then on quantum level corresponding to this size it is about 28.2 % of the whole universe energy. The rest of the universe energy 71.8 % occurs on near quantum levels.

If we are situated on quantum level at the size a from imaginary centre of our universe, we are able to observe only the mass situated on the same quantum level. It means that the rest of our universe mass is not observable for us even though it gravitationally influences our universe as a whole.

9　Cosmological shift of spectrum

Perception (measurement) of time flow was obviously different than it is today. Physical process lasting 1 s at present time lasted n_2/n_1 times longer in the past. Dimensions of mass objects were n_2/n_1 times bigger and photons radiated from them had n_2/n_1 times longer wavelength than they have in the same process today.

The shift of the spectrum of the radiation of the cosmological objects is defined:

$$z \equiv \frac{\lambda_r - \lambda_e}{\lambda_e} \,. \tag{53}$$

This relation presumes that the spectrum of cosmological source was the same in the past and today and the cosmological shift has happened during the travel from the source to an observer in consequence of the universe expansion. If the particles that create mass had smaller mass in the past than today then the energy radiated from them was equivalently smaller than today. We should rather write:

$$z = \frac{\lambda_r - \lambda_{e-today}}{\lambda_{e-today}} \,. \tag{54}$$

In case that the mass of elementary particles were smaller in the past, then:

$$\lambda_e = \lambda_{e-today} \frac{n_r}{n_e} \,. \tag{55}$$

According to (42), (54) and (55) results (as in classical theory):

$$z + 1 = \frac{\lambda_r}{\lambda_{e-today}} = \frac{\lambda_r}{\lambda_e} \frac{n_r}{n_e} = \frac{a_r}{a_e} \,. \tag{56}$$

10　Luminosity of cosmological sources

If the red shift does not exist, the apparent luminosity l of a cosmological source would be given by relation:

$$l = \frac{L}{S} \tag{57}$$

where:

L – absolute luminosity of a cosmological source

S – area on which photons from the cosmological source fall to.

The radiation energy from a cosmological source decreases in three ways:

1. The energy of the detected photons is lower then their original energy due to red shift according to (57).

2. Photons radiated during time interval $t_{e-today}$ (time the process would last today) will reach target during time interval Δt_r:

$$\frac{\Delta t_r}{\Delta t_{e-today}} = \frac{\lambda_r}{\lambda_{e-today}} = 1 + z. \tag{58}$$

3. We can't forget influence of lesser particle mass in the past:

$$\frac{\lambda_e}{\lambda_{e-today}} = \sqrt{\frac{a_r}{a_e}} = \sqrt{1 + z}. \tag{59}$$

The relative luminosity l of a typical cosmological source (cosmological candle) can be then written in the form:

$$l = \frac{L}{4\pi d_L^2} = \frac{L}{4\pi r_e^2 a^2 (1 + z)^{2.5}} \tag{60}$$

where:

d_L – distance of a cosmological source given by:

$$d_L = r_e a (1 + z)^{1.25}. \tag{61}$$

The variable r_e is given by [5] for $k = 1$ and $\ddot{a} = 0$ by the relation:

$$r_e = c \sin\left(\int_{t_e}^{t_r} \frac{dt}{a}\right) = \sin\left(\ln \frac{t_r}{t_e}\right) = \sin\left[\ln\left(1 + z\right)\right]. \tag{62}$$

The relative magnitude of stars m is defined by the Pogson equation [6]:

$$m = -2.5 \log\left(\frac{I}{I_0}\right) \tag{63}$$

where I_0 is the bolometric reference value $2.553 \times 10^{-8} \mathrm{W\,m^{-2}}$.

Now we can calculate value l (for suitable L) in the relation (60) and calculate the curve $m = m(z)$ using the relation (63) (see Fig. 2). The best fit with real measured values of relative magnitude of supernovas type Ia [7] we get for $L \cong 2.765 \times 10^{28}$ W. It acknowledges that the model above can correspond with our reality.
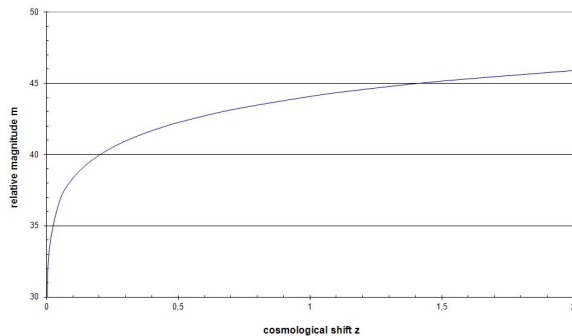

Fig. 2: Relative supernova magnitude – calculated for $L = 2.765 \times 10^{28}$ W

We can construct the so-called residual Hubble diagram – relative luminosity of supernovas related to the case of an empty universe ($\Omega = 0, k = -1, q = 0$) (see Fig. 3).

$$\Delta(m - M) = 5 \log\left(\frac{r_e}{r_{e0}}\right). \tag{64}$$

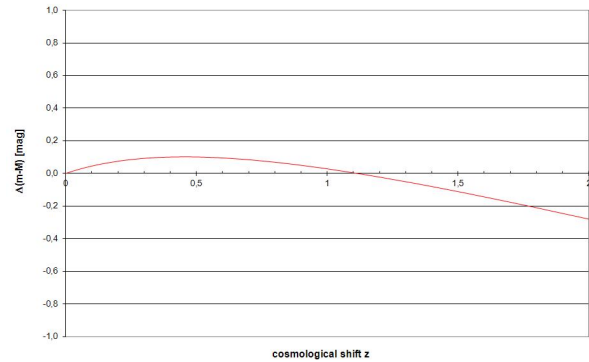$$r_{e0} = \sinh[\ln(1 + z)]. \tag{65}$$

Fig. 3: Residual Hubble diagram – without consideration of dust influence

11 Conclusion

Our universe doesn't have to be necessarily open and accelarating its expansion in order to be in accordance with our present observation and knowledge. In this article, I tried to show that our universe can be closed and uniformly expanding supposing that its mass increases proportionally to its size and analogically its size increases proportionally to its mass, similarly as black holes do.

References

1. Bohm D., Factor D., ed. Unfolding Meaning: A weekend of dialogue with David Bohm. Foundation House, Gloucestershire, 1985.

2. Pathria R. K. The universe as a black hole. *Nature*, 1972, v. 240 (5379), 298–299.

3. Skalský V. The model of a flat (Euclidean) expansive homogeneous and isotropic relativistic universe in the light of general relativity, quantum mechanics, and observations. *Astrophys. Space Sci.*, 2010, v. 330, 373–398.

4. Cely J. Zaklady kvantove mechaniky pro chemiky (Fundamentals of quantum mechanics for chemists). Univerzita Jana Evangelisty Purkyne, Bmo, 1986, p. 45.

5. Misner C. W., Thorne K. S., Wheeler J. A. Gravitation. W. H. Freeman and Co., San Francisco, USA, 1973.

6. http://aldebaran.cz/studium/astrofyzika.pdf, p. 7.

7. http://www.astro.ucla.edu/~wright/sne_cosmology.html [18.4.2008].

8. http://aldebaran.cz/bulletin/2004_33_dma.html [18.4.2008].

Approach to the Schwarzschild Metric with SL(2,R) Group Decomposition

Alexander Kritov

E-mail: alex@kritov.ru

The paper analyzes the two-step coordinate transformations, known as the simple (or "heuristic") approach to the Schwarzschild metric [3, 5, 22]. The main finding of the analysis is that such transformations are *unique* as they correspond to the Iwasawa decomposition for the special linear group $SL(2, \mathbb{R})$ with the subgroup of rotation $SO(1, 1)^+$. It is noted that all original transformations utilize de facto determinant of unity. However, as shown, this property is related to the action invariance under diffeomorphism for gravity. The noted group symmetry of the coordinate transformations may shed light on the "paradox" of the original approach for obtaining the Schwarzschild metric based on the Equivalence Principle only and enable its further study. The path to generalization in $SL(4, \mathbb{R})$ is suggested.

1 Introduction

In work "What is wrong with the Schwarzschild coordinates" [5], J. Czerniawski demonstrated the two-step coordinate transformations from the Minkowski to the Schwarzschild metric. Recently, Christillin and Morchio [3] slightly updated the approach by clarifying the step of the transformation from the Gullstrand-Painlevé (G-P) to the Schwarzschild metric. Without this, the original path would not be consistent. Even if the approach of obtaining the Schwarzschild metric via the "heuristic" to be considered with certain cautiousness, the original work was over-cited, bringing the substantial interest in this topic [1, 3, 7, 14, 22]. The approach recently was called the "inherent paradox of GR" [3], and the original question has not been answered. This paper aims to walk through the approach with maximum formality to present the correspondence and possible path to the generalization.

2 Preliminaries and Notation

Diffeomorphism of a manifold $\mathcal{M}$ by definition is a smooth invertible map $\phi : \mathcal{M} \to \mathcal{M}$ such as the inverse map ϕ^{-1} be smooth as well. General diffeomorphism can be thought as the deformation that does not preserve the metric on $\mathcal{M}$. The map $\phi : M \to M$ of the transformation from affine η to curvilinear g coordinates may be considered as a vector-valued function of n-variables. By retaining the requirements of smoothness, the transformation may be defined in terms of partial derivatives in the form of the Jacobian matrices that constitute second rank tensors

$$J_{\mu a} = \frac{\partial x^a}{\partial \bar{x}^\mu} \quad J^{\mu a} = \frac{\partial \bar{x}^a}{\partial x^\mu} . \tag{1}$$

The barred symbols denote the curvilinear coordinates, and unbarred are for flat coordinates*. The metric tensor is

$$g_{\mu\nu} = J_{\mu a} J_{\nu b} \, \eta_{ab} \quad g^{\mu\nu} = J^{a\mu} J^{b\nu} \eta^{ab} \tag{2}$$

---

*Since the order of indexes for J in the notation is arbitrary, it is chosen such as the covariant form coincides with the "vierbein" or tetrad. So, one can treat them as the same objects.

where indexes are $(0, 1, 2, 3)$ and η has the signature $(- + + +)$. The transformation is non-singular $J \neq 0$, the matrix is bijective, and the inverse transform represents the simple inverse matrix $\bar{J} = J^{-1}$. If the order of indexes as per (1), the equation can be written in the matrix notation (for both covariant and contravariant forms) as

$$g = J \cdot \eta \cdot J^T . \tag{3}$$

The capital letters are used for matrices excluding the metric tensor g, and Minkowski η. In matrix notation, the form (covariant or contravariant) will be specified in the text. For the spherical symmetry case, the Jacobian matrices are 4×4

$$\begin{bmatrix} J & 0 \\ 0 & I_2 \end{bmatrix} .$$

Therefore, J can be written as 2×2 for the temporal and radial coordinates only, dropping the symmetric angular and tangential terms that are not affected by transformations. The spherical symmetry provides the unique case to consider the transformations as being "two-dimensional" with certain limitations. Though later in Section 7, the four-dimensional form is reviewed. Natural units $(c = 1)$ are employed throughout. As a matter of choice, the common hyperbolic notation is used for the radial escape velocity for shortness

$$v = \text{th}(\beta) = \sqrt{\frac{r_g}{r}}$$

$$\sinh(\beta) = \frac{v}{\sqrt{1 - v^2}}$$

$$\gamma = \cosh(\beta) = \frac{1}{\sqrt{1 - v^2}} = \frac{1}{\sqrt{1 - \frac{r_g}{r}}}$$

3 Step one: from Minkowski to Gulfstrand-Painlevé

The first coordinate transformation as given in [5, 22] is

$$dx^1 = d\bar{x}^1 - v d\bar{x}^0 \qquad dx^0 = d\bar{x}^0 \tag{4}$$

where v is the radial escape velocity of the gravitational field or the river velocity [7]. The equations have the differential form; therefore, the term "Galilean transformations" can be used with certain cautiousness. Despite the similarity in the look, the latter is defined as the affine transformations of the coordinates[*]. According to [3, 5], this transformation embodies the Equivalence Principle (EP) and therefore plays the central role in the approach.

The Jacobian matrix for the first transformation as per definition (1) is then

$$J_{\mu a}^{(1)} = \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \text{th}(\beta) \\ 0 & 1 \end{bmatrix}$$
$$J_{(1)}^{\mu a} = \begin{bmatrix} 1 & 0 \\ v & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \text{th}(\beta) & 0 \end{bmatrix} \quad (5)$$

where v can be taken with an arbitrary $\pm$ sign as not affecting the final transform [5, 14]. Such transformation can be classified as the spacetime shear deformation. It obviously represents shear mapping transformation on the hyperbolic plane. The value of "shear" is given by the relativistic velocity $v = \text{th}(\beta)$ and the (imaginary) shear angle is β or rapidity. Further, the term shear is used for this transformation for the current purposes leaving aside its physical significance and the relation to the EP. It leads to the G-P coordinates with the metric tensor which has following covariant form

$$g_1 = J^{(1)} \cdot \eta \cdot J^{(1)T} = \begin{bmatrix} -\left(1 - v^2\right) & v \\ v & 1 \end{bmatrix}$$
$$= \begin{bmatrix} -\left(1 - \frac{r_g}{r}\right) & \sqrt{\frac{r_g}{r}} \\ \sqrt{\frac{r_g}{r}} & 1 \end{bmatrix} . \quad (6)$$

4   Step two: to the Schwarzschild metric

The second coordinate transformation J_2 is pull-back from the comoving G-P frame to the coordinate frame of reference redefining time coordinate. The covariant form is

$$J^{(2)} = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \quad (7)$$

where b is the *arbitrary* parameter [†]. The total coordinate transformation is the product of both transforms

$$J = J^{(2)} \cdot J^{(1)} = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & v \\ b & vb + 1 \end{bmatrix} \quad (8)$$

---

[*]The differential form of the Lorentz transformations has the same form and obviously $\Lambda \eta \Lambda^T = \eta$ is valid for the differential form. For more on differential transformation see [8].

[†]As suggested in [3] "the requirement to eliminate the off-diagonal term of the P-G metric is generally accomplished just by redefining time in an ad hoc way".

that leads to the metric tensor

$$g = J \cdot \eta \cdot J^T = \begin{bmatrix} -(1 - v^2) & (v^2 - 1)\,b + v \\ (v^2 - 1)\,b + v & (vb + 1)^2 - b^2 \end{bmatrix} . \quad (9)$$

Choosing b in the way to eliminate the off-diagonal terms one obtains the Schwarzschild metric

$$g_{\mu\nu} = \begin{bmatrix} -\left(1 - v^2\right) & 0 \\ 0 & \left(1 - v^2\right)^{-1} \end{bmatrix}$$
$$= \begin{bmatrix} -\cosh^{-2}(\beta) & 0 \\ 0 & \cosh(\beta) \end{bmatrix} . \quad (10)$$

After b has been defined, the second transformation becomes

$$J_{\mu a}^{(2)} = \begin{bmatrix} 1 & 0 \\ \sinh(\beta)\cosh(\beta) & 1 \end{bmatrix}$$
$$J_{(2)}^{\mu a} = \begin{bmatrix} 1 & \sinh(\beta)\cosh(\beta) \\ 0 & 1 \end{bmatrix} . \quad (11)$$

As a result, the parameter $\pm b = v\gamma^2 = \sinh(\beta)\cosh$ corresponds to the proper velocity of free-falling observer in the Schwarzschild metric. It stands in the well-known expression for the time coordinate transformation between the G-P and the Schwarzschild metrics.

5 $SL(2, \mathbb{R})$ with the Lorentz signature

The remarkable property of all Jacobian matrices is that they all have the unity determinant[‡]. In order to classify them as elements of a group, one may note that matrices are defined on the Minkowski basis (space-time or the hyperbolic plane). In fact, the Jacobian matrices can be expressed using an imaginary value for the time coordinate as

$$J_{\mu a} = \frac{\partial x^a}{\partial \overline{x}^\mu} = \begin{bmatrix} \frac{\partial x_0}{\partial \overline{x}_0} & \frac{1}{i}\frac{\partial x_1}{\partial \overline{x}_0} \\ i\,\frac{\partial x_0}{\partial \overline{x}_1} & \frac{\partial x_1}{\partial \overline{x}_1} \end{bmatrix} . \quad (12)$$

In such a way, the Jacobian matrices constitute the subgroup of SL(2, $\mathbb{C}$) with only two imaginary off-diagonal elements in the matrices. Let's denote this group as SL(2, $\mathbb{C}$)* $\in$ SL(2, $\mathbb{C}$). Then, considering only the real parts, there is one-to-one mapping of $Z' \in$ SL(2, $\mathbb{C}$)* to $Z \in$ SL(2, $\mathbb{R}$) as follows

$$Z' = \begin{bmatrix} a & -i\,b \\ i\,c & d \end{bmatrix} \to Z = \begin{bmatrix} a & b \\ c & d \end{bmatrix} . \quad (13)$$

Ignoring the imaginary unit, in the way as it is done for the Minkowski time coordinate, allows one to use the real values in the matrix as per the defined mapping to SL(2, $\mathbb{R}$). Introduced in such a way, the group SL(2, $\mathbb{C}$)* is isomorphic to SL(2, $\mathbb{R}$). This mapping is multiplicative and a bijection. Hence, all operations in SL(2, $\mathbb{R}$) can be translated to SL(2, $\mathbb{C}$)* and vice versa using this isomorphism. Such mapping allows one to utilize SL(2, $\mathbb{R}$) on the Lorentz/Minkowski basis $\mathbb{H}^{1(2)}$, instead of its default, the Euclidean basis $\mathbb{R}^2$.

[‡]To be consistent, the fact is taken *a priori* "knowing" that the resulting metric has $|g| = |\eta| = -1$. Section 8 reviews a physical ground for this.

6　The group decomposition

The Iwasawa decomposition is the factorization of a semisimple Lie group to the product of three closed subgroups as K × A × N ("compact, Abelian and nilpotent") [9, 13]. In the application to SL(2, $\mathbb{R}$) it is well studied [4, 12], and in terms of the matrices is even obvious. Importantly, it implies the *uniqueness* of the factorization of the element of the group to the product of three subgroups, those elements are N is upper triangular, A is diagonal, and K is orthogonal matrices, the spatial rotations K ∈ SO(2).

One may see that using the mapping (13), the elements of these three groups become the matrices of the following form

$$N = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \quad A = \begin{bmatrix} k & 0 \\ 0 & k^{-1} \end{bmatrix}$$
$$K = \begin{bmatrix} \cosh(\alpha) & \sinh(\alpha) \\ \sinh(\alpha) & \cosh(\alpha) \end{bmatrix} \tag{14}$$

with $k > 0$. Since the mapping results in the complex conjugation of the angle of rotation ($\beta \to i\beta$), the foremost notable distinction from the decomposition of SL(2, $\mathbb{R}$) is that K becomes the group of hyperbolic rotations SO(1, 1)$^+$, that is the pure Lorentz boost.

The covariant form of $J^{(1)}$, and contarvariant $J_{(2)}$ are ∈ N (upper triangular matrices). Therefore, the decomposition can be applied to contravariant $J_{(1)}$ and to contravariant $J^{(2)}$ which are lower triangular. In fact, they are explicitly the Iwasawa decomposition $J^{(2)} = A \cdot K \cdot N$ (covariant form) and $J_{(1)} = N \cdot A \cdot K$ (contravariant form). The latter is as follows

$$J_{(1)}^{\mu a} = \begin{bmatrix} 1 & 0 \\ \mathrm{th}(\beta) & 0 \end{bmatrix} = \begin{bmatrix} 1 & -\sinh(\beta)\cosh(\beta) \\ 0 & 1 \end{bmatrix} \cdot$$
$$\cdot \begin{bmatrix} \cosh(\beta) & 0 \\ 0 & \cosh^{-1}(\beta) \end{bmatrix} \cdot \begin{bmatrix} \cosh(\beta) & \sinh(\beta) \\ \sinh(\beta) & \cosh(\beta) \end{bmatrix} . \tag{15}$$

Notably, that N in the factorization becomes already known matrix $N = J_{(2)}^{-1}$ (11). The resulting transformation is

$$J = J_{(2)} J_{(1)} = A \cdot K \tag{16}$$

where $J_{(2)} J_{(1)}$ has the form of the product of two upper and lower triangular matrices $N_1 \cdot \overline{N}_2$. And since K ≡ Λ is the Lorentz boost, that leaves the original metric invariant $\eta = \Lambda \cdot \eta \cdot \Lambda^T$, then K drops being at the right side of (16). Therefore the resulting Schwarzschild metric

$$g = J \cdot \eta \cdot J^T = A \cdot \eta \cdot A^T \tag{17}$$

is obviously defined by the diagonal matrix A*

$$A^{vb} = \begin{bmatrix} \cosh(\beta) & 0 \\ 0 & \cosh^{-1}(\beta) \end{bmatrix}$$

$$A_{vb} = \begin{bmatrix} \cosh^{-1}(\beta) & 0 \\ 0 & \cosh(\beta) \end{bmatrix} . \tag{18}$$

Therefore, all approach can be represented as just the diagonalization of the first shear transformation matrix.

Proposition: If J_1 is the shear transformation in the contravariant form with the shear value v, then its Iwasawa decomposition with the mapping (13) provides the diagonal matrix A that uniquely represents the Jacobian matrix J that maps the Minkowski to the Schwarzschild metric. The process is that A normalizes N, or A is the unique diagonal form of the original shear transformation[†].

7　The generalization to the Cartesian coordinates

The suggested approach can be generalized to four-dimensional spacetime in the Cartesian coordinates. The hyperbolic shear parameter v is non-Lorentz invariant four-vector $v = (1, v_x, v_y, v_z)$, and its norm is $\|v\| = \cosh(\beta)^{-1}$. It shall constitute the column of contravariant shear transformation in the Cartesian coordinates[‡]

$$J_{(1)}^{va} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ v_x & 1 & 0 & 0 \\ v_y & 0 & 1 & 0 \\ v_z & 0 & 0 & 1 \end{bmatrix} . \tag{19}$$

The KAN decomposition of this form provides the unique Jacobian matrix for the metric as described in the Proposition. In case if $v_y = v_z = 0$, implying that one direction via coordinate x is considered, then it converges to the reviewed case above. It is known that the Iwasawa decomposition can be also applied to elements of SL(4, $\mathbb{R}$) group [4, 19]. The straightforward approach is to use the Gram–Schmidt process that leads to QR decomposition, from which the KAN form can be obtained [19]. However, the more elegant way is to use the Givens rotations, which are literally spatial rotations of the SO(3) group. Obviously, the shear vector in the Cartesian coordinates can be represented as

$$v = (1, \mathrm{th}(\beta)\sin(\theta)\cos(\phi), \mathrm{th}(\beta)\sin(\theta)\sin(\phi), \mathrm{th}(\beta)\cos(\theta))$$

where θ and ϕ are the angles between vector v and the coordinate axes. Hence, two sequential spatial rotations $R_z(\phi) \in$ SO(3) and $R_y(\frac{\pi}{2} - \theta) \in$ SO(3) reduce the matrix to the case above, eliminating second and third components (v_y and v_z). Treated in such way, a general transformation in four-dimensional spacetime (19) is {SL(2, $\mathbb{R}$), SO(3)}.

The details and the analysis of the decomposition of (19) lay out of the scope of this work and can be an interesting topic for future research.

---

*It coincidences with the Schwarzschildian vierbein or "metric squared".

[†]NAK, as shown, results in contrvariant form of A, similarly KAN decomposition gives the covariant form of A.

[‡]Note, the Jacobian's column vectors' signature becomes *opposite* to the metrics signature (η and g) as per definition of SL(2, C)* above.

8 Discussion

At the critical angle, a possible weak point of the original path should be also noted. When one uses "one coordinate change" transformations (5) and (7), in fact, the additional condition on the determinant $|J| = 1$ is taken "under the hood". During the classical derivation of the Schwarzschild metric in the GR, $|g| = -1$ is the obtained results from the field equations (note: even with $T_{\mu\nu} = 0$). Contrary to that, the reviewed "heuristic" approach uses $|J| = 1$ that explicitly leads to $|g| = -1$ *a priori* knowing the resulting metric.

Once this principle physically has solid ground, then the above parallel can be considered fundamental. Without this, one may still regard this approach as a coincidence. From the prospect of the physics, the value of g_{00} for the Schwarzschild metric can be obtained from the Newtonian gravitation [15] or the equivalence principle and red-shift experiments [20, 21]. If one would a priori know that $|g| = -1$, then the Schwarzschild metric easily follows by defining $g_{rr} = -g_{00}^{-1}$.

From another perspective, the fact is that the spherically symmetric static gravitational field has explicitly $|g| = -1$ cannot be just a coincidence but may potentially signal a hidden symmetry attached to such property.

Consider the action in the Minkowski spacetime $S_1(x) = \int \mathcal{L}(x, \dot{x}) \, dV^4$ and in the spacetime with the curvature $S_2(x) = \int \sqrt{-g} \, \mathcal{L}(x, \dot{x}) \, dV^4$ expressed by the Lagrangian density. The diffeomorphism invariance of the action would require that under the map $\phi : S_1 \to S_2 = S_1$ and therefore $|g| = -1$. On the other hand, the action invariance under diffeomorphism implies the equivalence of the conservation of energy, momentum, and the continuity equations for the system.

9 The conclusion

The analyzed approach shows the striking correspondence between coordinate transformation from the Minkowski spacetime to the Schwarzschild metric and SL(2, $\mathbb{R}$) group using the mapping to the Lorentz base. The original "heuristic" approach to the Schwarzschild metric can be considered via the unique group decomposition by obtaining the first coordinate transformation's corresponding diagonal form.

SL(2, $\mathbb{R}$) group has already appeared in the application to the gravitation metric in [10] and in two-dimensional quantum gravity [17]. This review gives a more classical and intuitive outlook on the group's correspondence to the coordinate transformations of the metrics.

The work outlines a critical point of the original approach, though suggesting further prospects for the method generalization and research. The reviewed case brings an additional question on the action invariance under diffeomorphism for the gravity. The group symmetry of the reviewed coordinate transformations may probably shed light on the resolution of the mentioned "inherent paradox of GR".

References

1. Barceló C., Liberati S., Visser M. Analogue Gravity. *Living Rev. Relativ.*, 2011, v. 14, 3. arXiv: gr-qc/0505065.

2. Bourbaki N. Lie Groups and Lie Algebras. Springer-Verlag, Berlin, 1989.

3. Christillin P., Morchio G. Relativistic Newtonian gravitation. arXiv: gr-qc/1707.05187.

4. Conrad K. Decomposing SL2(R). https://kconrad.math.uconn.edu/blurbs/, accessed 2020.

5. Czerniawski J. What is wrong with Schwarzschild's coordinates? arXiv: gr-qc/0201037.

6. Gilmore R. Relations Among Low-Dimensional Simple Lie Groups. *Geometry and Symmetry in Physics*, 2012, v. 28, 1–45.

7. Hamilton A. J. S., Lisle J. P. The river model of black holes. *American Journal of Physics* 2008, v. 76, 519–532. arXiv: gr-qc/0411060.

8. Huang, Y.-S. A new perspective on relativistic transformation: formulation of the differential Lorentz transformation based on first principles. *Phys. Scr.*, 2010, v. 82, 045011.

9. Husemöller D., Joachim M., Jurčo B., Schottenloher M. Gram–Schmidt Process, Iwasawa Decomposition, and Reduction of Structure in Principal Bundles. In: Basic Bundle Theory and K-Cohomology Invariants. Lecture Notes in Physics, vol. 726, Springer, Berlin, 2008.

10. Jadczyk, A. Gravitation on a Homogeneous Domain. arXiv: math-ph/1105.3814v1.

11. Kassner K. A physics-first approach to the Schwarzschild metric. *Advanced Studies in Theoretical Physics*, 2017, v. 11 (4), 179–212. arXiv: gr-qc/1602.08309.

12. Kisil, Vladimir V. Geometry of Möbius Transformations: Elliptic, Parabolic and Hyperbolic Actions of SL2. World Scientific, 2012.

13. Knapp, A. W. Lie Groups Beyond an Introduction. Birkhäuser, Basel, 2002.

14. Kritov A. Unified Two Dimensional Spacetime for the River Model of Gravity and Cosmology. *Progress in Physics*, 2019, v. 15 (3), 163–170.

15. Landau L. D., Lifshitz E. M. The Classical Theory of Fields. Butterworth-Heinemann, 1987.

16. Moretti V., Pinamonti N. Holography and SL(2, R) symmetry in 2D Rindler space-time. *Journal of Mathematical Physics*, 2004, v. 45, 230.

17. Polyakov A. M. Quantum Gravity in two dimensions. *Modern Physics Letters A*, 1987, v. 2 (11), 893–898.

18. Sacks W. M., Ball J. A. Simple Derivation of the Schwarzschild metric. *American Journal of Physics*, 1968, v. 36, 240.

19. Sawyer P. Computing the Iwasawa decomposition of the classical Lie groups of noncompact type using the QR decomposition. *Linear Algebra and its Applications*, 2016, v. 493, 573–579.

20. Schild A. Equivalence Principle and Red-Shift Measurement. *American Journal of Physics*, 1960, v. 28, 778.

21. Shiff L. I. On Experimental Tests of the General Theory of Relativity. *American Journal of Physics*, 1960, v. 28, 340.

22. Visser M. Heuristic Approach to the Schwarzschild geometry. arXiv: gr-qc/0309072.

A Wave Representation for Massless Neutrino Oscillations: The Weak Interaction Transmutes the Wave Function

Edward R. Floyd

10 Jamaica Village Road, Coronado, California 92118, USA. E-mail: floyd@san.rr.com

There are solutions of the Klein-Gordon equation for the massless neutrino that produce massless neutrino oscillation of flavor. These solutions serve as a counterexample to Pontecorvo, Maki, Nakagawa, and Sakata theory for neutrino oscillation of flavor, which implies neutrinos must have mass contrary to the standard model. We show that the wave function for the massless antineutrino for an inverse β decay (IBD) is a superposition of two independent solutions of the Klein-Gordon equation. One solution represents the latent incident wave upon an IBD. The other solution represents the latent reflected wave from the IBD. This superposition renders a compound modulated wave function with regard to amplitude and phase modulations. This compound modulation is shown to facilitate neutrino oscillation that may be massless and, therefore, consistent with the standard model. Extra to a massless counterexample, the weak interaction is shown to transmute the wave function during an IBD by changing the amounts of the latent incident and latent reflected wave functions that are allocated to the superposition.

1 Introduction

The Pontecorvo, Maki, Nakagawa, and Sakata (PMNS) theory for oscillation of neutrino (ν) flavor implies that the neutrino has a finite mass in contrast to the standard model [1]–[4]. PMNS theory, which was developed in the mid-twentieth century in the absence of a contending theory, soon became preeminent regarding neutrino oscillations including its implication that the neutrino must have a finite mass in order to oscillate. A counterexample to PMNS theory now exists: the quantum trajectory representation of quantum mechanics had predicted in 2017 that massless neutrino oscillation is an alternative possibility that is consistent with the standard model [5]. However, the quantum trajectory representation is presently arcane, for it is couched in a quantum Hamilton-Jacobi formulation [5]–[17]. As a result, PMNS theory has maintained its preeminence on neutrino oscillation. A way to overcome this preeminence is to describe **massless** neutrino oscillation in the more familiar wave function representation, which would be more accessible to a much broader audience. Our objective in this paper is to provide such.

A wave function representation that is a counterexample to PMNS theory is attainable. This theoretical counterexample renders massless neutrino oscillation while also showing that PMNS theory is not the exclusive explanation of neutrino oscillation. In this paper, we show that there are mathematical solutions of wave equations, which to the best of our knowledge have been used only a few times [18]–[23] to describe wave phenomena, and which invite further investigation. We study massless neutrino oscillation with these mathematical solutions of the Klein-Gordon equation for a massless antineutrino. This mathematical solution is synthesized by the superpositional principle from two independent solutions of the Klein-Gordon equation for an antineutrino before

encountering a charged current interaction. The two solutions are the latent incident solution and the latent reflected solution. The "quantum action" of the Klein-Gordon equation is composed of both independent solutions of the Klein-Gordon equation [14] and can be seen as the order $\hbar^0$ term of the quantum action of QFT.

Extra to the initial goal of adducing a massless counterexample, the behavior of the synthesized solution also gives insight into the weak interaction (weak force). A byproduct of this investigation shows that the weak interaction without causing any exchange of energy can transmute the Klein-Gordon solution from a synthesized solution to a plane-wave solution.

The particular charged current interaction that we examine herein is the inverse beta decay (IBD) where [24]

$$\overline{\nu}_e \;+\; \mathrm{p} \;\xrightarrow{\;\;\mathrm{W^+\;boson\;exchange}\;\;}\; \mathrm{e^+} \;+\; \mathrm{n}, \qquad (1)$$
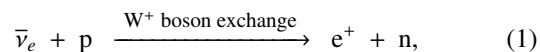
in which the antineutrino $\overline{\nu}$ participates as an electron antineutrino $\overline{\nu}_e$. The wave function for $\overline{\nu}$ is specified by $\overline{\psi}$. When $\overline{\nu}_e$ arrives at the point q_b ready for IBD absorbtion in (1), its $\overline{\psi}$ is assumed in this ab initio calculation to be then a traveling complex-exponential plane wave $\exp(ikq)$ with wave number k, in cartesian coordinate q, and tacitly with amplitude 1. While the ab initio calculation develops flavor oscillations for a massless $\overline{\nu}$, the conventional terminology "neutrino oscillation" is retained for referencing the oscillation phenomenon herein.

An outline of the rest of this paper follows. In §2 we develop a model by an ab initio computation for massless neutrino oscillation for an IBD. The wave function for the neutrino is synthesized from the latent solutions for the incident and reflected wave functions by the superpositional principle. The latent incident and latent reflected wave functions are

traveling complex-exponential plane waves that are independent one-dimensional solutions of the Klein-Gordon equation. This synthesized solution is shown to be compoundly modulated with regard to amplitude and phase. This compound modulation induces periodic nonuniform propagation that in turn facilitates neutrino oscillation. The amplitude and phase modulations are individually analyzed. We apply the same modulation analyses to the wave function's spatial derivative. In this wave function representation for massless oscillation, the weak interaction changes the synthesized wave function to a traveling complex-exponential plane-wave solution, which is then ready for absorption by the IBD process. In §3, we examine selected didactic examples. The examples show that the individual contributions of phase modulation and amplitude modulation complement each other. Where one modulation is at a peak, the other is at a null. The examples also show that the compound modulations of the wave function and its derivative supplement each other. That is where the amplitude modulation increases dilation in one, it decreases it in the other. And where phase modulation rotates the phase of one clockwise, it rotates the other's phase counterclockwise. In §4 a brief discussion is presented. Together, the complementing and supplementing are shown to facilitate periodic nonuniform propagation that permits massless neutrino oscillation. Findings and conclusions are presented in §5.

2 Ab initio calculation

The one-dimensional stationary Klein-Gordon equation (SKGE) for an antineutrino with mass m and for the Cartesian dimension q is a second-order, linear, homogeneous ordinary differential equation given by [25]

$$-\hbar^2 c^2 \frac{\partial^2 \overline{\psi}(q)}{\partial q^2} + \left(m^2 c^4 - E^2\right) \overline{\psi}(q) = 0 \qquad (2)$$

where $\hbar$ is Plank's constant, c is speed of light and E is energy. As such, the superpositional principle applies to the SKGE's solutions. The inertial reference frame for describing $\overline{\psi}$ of (2) is the frame for which the target proton of the IBD is at rest. This makes E dependent on the dynamics of the target proton. The threshold energy for executing an IBD is $E_{\text{threshold}} = 1.806$ MeV for ν_e and progressively greater for the analogous charged current interactions for ν_μ and ν_τ. Herein, it is always assumed the $\overline{\nu}$ has energy greater than the threshold energy. The notation $\overline{\psi}$ denotes that the wave function of the antineutrino is a solution of (2) but does not specify whether it is unispectral, $\overline{\psi} = \exp(ikq)$, or bispectral $\overline{\psi}_2$. Eq. (2) remains well posed should $m = 0$ in agreement with the standard model. Studying the case $m = 0$ is sufficient to render a massless counterexample to PMNS. For antineutrino energy E and nil mass, a set of independent solutions sufficient to solve (2) may be given by $\{\overline{\psi}, \breve{\psi}\} = \{\exp(+ikq), \exp(-ikq)\}$ where the wave number $k = E/(\hbar c)$.

The incident antineutrino is assumed to propagate in the $+q$ direction toward the target proton of an IBD, while any reflection from an IBD would propagate in the $-q$ direction. The solution $\overline{\psi} = \exp(ikq)$ is a unispectral wave function with one spectral component, $+k$ (the solution of the homogeneous SKGE is defined to within a constant in phase). Its derivative $\partial_q \overline{\psi} = ik\overline{\psi}$ is also unispectral and is displaced in phase from $\overline{\psi}_1$ by a constant $\pi/2$ radians. The amplitude of $\partial_q \overline{\psi}$ relative to that of $\overline{\psi}$ is multiplied by the factor k. Thus, the unispectral $\overline{\psi}(q)$ displays uniform rectilinear motion, which presents a constant relationship

$$\partial_q \overline{\psi} / \overline{\psi} = \partial_q \ln(\overline{\psi}) = ik \qquad (3)$$

to any encountered current interactions. The constant character of (3) is expected, for $\overline{\psi}(q)$ is an exponential of the linear variable q. Uniform rectilinear propagation precludes flavor oscillations.

Let the incident antineutrino to an IBD have a bispectral wave function $\overline{\psi}_2$ with spectral components given by wave numbers $\{+k, -k\}$. We can synthesize a bispectral $\overline{\psi}_2$ by the superpositional principal from the set $\{\exp(+ikq), \exp(-ikq)\}$ of independent solutions for the SKGE. The incident bispectral $\overline{\psi}_2$ may be presented in a few representative forms as [5]

$$\overline{\psi}_2 = \overbrace{\underbrace{\alpha \exp(+ikq)}_{\text{latent incident wave}} + \underbrace{\beta \exp(-ikq)}_{\text{latent reflected wave}}}^{\text{bispectral solution of SKGE by superpositional principle}} \qquad (4)$$

$$= \underbrace{(\alpha - \beta)\exp(ikq)}_{\text{traveling wave}} + \underbrace{2\beta \cos(kq)}_{\text{standing wave}}$$

$$= \underbrace{(\alpha + \beta)\cos(kq) + i(\alpha - \beta)\sin(kq)}_{\text{coherent standing waves}} \qquad (5)$$

$$= \underbrace{A_{\overline{\psi}} \exp(i\, P_{\overline{\psi}}),}_{\text{compoundly modulated traveling wave}} \qquad (6)$$

where all forms (4)–(6) are solutions of the SKGE. In (6), $\overline{\psi}_2$ is compoundly modulated for its amplitude $A_{\overline{\psi}}$ and phase $P_{\overline{\psi}}$ are modulated as given by

$$A_{\overline{\psi}} = \overbrace{\left[\alpha^2 + \beta^2 + 2\alpha\beta \cos(2kq)\right]^{1/2}}^{\text{amplitude modulation}}$$

and

$$P_{\overline{\psi}} = \overbrace{\arctan\left(\frac{\alpha - \beta}{\alpha + \beta}\tan(kq)\right)}^{\text{phase modulation [5]}}.$$

Eqs. (4)–(6) for the antineutrino's wave function are all representations of a wave function synthesized by the superpositional principle. As such, each individual equation of (4) through (6) represents a synthesized solution of the SKGE consistent with the orthodox interpretation of quantum mechanics. The coefficients α and β respectively specify the

amplitudes for the latent incident and reflected waves associated with an IBD. Propagation of the latent incident wave in the $+q$ direction implies that $\alpha^2 > \beta^2$. The coefficients α and β are normalized by

$$\alpha^2 - \beta^2 = 1 \qquad (7)$$

consistent with one $\overline{\nu}_e$ in (1) for an IBD (it is also the normalization used in the quantum trajectory representation). Knowing the value of one coefficient implies knowing the value of the other by normalization, (7). If the conditions $\alpha > 1$ and $0 < \beta^2 = \alpha^2 - 1$ exist, then bispectral propagation in the $+q$ direction follows. The bispectral propagation for $\overline{\nu}$ consistent with (4)–(6) is nonuniform, albeit still rectilinear, in the $+q$ direction. As such, $\overline{\psi}_2(q)$ may also be considered to be the wave function synthesized by the superposition of the latent incident wave and the the latent reflected wave upon each other. Note that herein the coefficients could have been expressed hyperbolically by $\alpha = \cosh(\gamma)$ and $\beta = \sinh(\gamma)$ consistent with (7).

For completeness, if the incident and reflected waves were neither latent nor superimposed, then the wave function representation would be in a two-dimensional space $\{q_{\text{incident}}, q_{\text{reflected}}\}$ given by

$$\overline{\psi}(q_{\text{incident}}, q_{\text{reflected}}) = \alpha \exp(+ikq_{\text{incident}})$$
$$+ \beta \exp(-ikq_{\text{reflected}}),$$

which is not equivalent to $\overline{\psi}_2(q)$ of (4)–(6). Eqs. (4)–(6) individually show the superpositioning to describe $\overline{\psi}_{\text{superimposed}}$ in one-dimensional space by a single independent variable q. Also for completeness, a literature search for "reflected neutrinos" on the web has found nothing for reflected neutrinos from charged current interactions *per se* but did find an unpublished report regarding reflections of antique neutrinos from the big bang [26].

Let us examine the compoundly modulated traveling wave (6) in special situations for didactic reasons. Should $\beta = 0$, then the amplitude $A_{\overline{\psi}}$ and phase $P_{\overline{\psi}}$ would respectively become

$$A_{\overline{\psi}}\big|_{\beta=0} = [\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)]^{1/2}\big|_{\beta=0} = \alpha|_{\beta=0} = 1 \quad (8)$$

and

$$P_{\overline{\psi}}\big|_{\beta=0} = \arctan\left(\frac{\alpha - \beta}{\alpha + \beta}\tan(kq)\right)\bigg|_{\beta=0} = kq. \qquad (9)$$

Then, (6) would represent unispectral propagation as expected. Next, we consider the case $(|\beta| = \alpha) \notin \{0 \le \beta^2 = \alpha^2 - 1\}$ and in violation of the normalization (7). Nevertheless, $|\beta| = \alpha$ is a limit point for $\beta \to \infty$. Should $\pm\beta = \infty$ (*i.e.* where a latent total reflection would preempt any IBD), then the amplitude would reduce to trigonometric identities with scaling factor 2α given by [27]

$$A_{\overline{\psi}}\big|_{\beta=\alpha} = 2\alpha\left(\frac{1 + \cos(2kq)}{2}\right)^{1/2} = 2\alpha\cos(kq) \qquad (10)$$

and

$$A_{\overline{\psi}}\big|_{-\beta=\alpha} = 2\alpha\left(\frac{1 - \cos(2kq)}{2}\right)^{1/2} = 2\alpha\sin(kq) \qquad (11)$$

consistent with (5). The corresponding phase would be

$$P_{\overline{\psi}}\big|_{\beta=\alpha} = \arctan\left(\frac{\alpha - \beta}{\alpha + \beta}\tan(kq)\right)\bigg|_{\beta=\alpha} = 0 \qquad (12)$$

and

$$P_{\overline{\psi}}|_{-\beta=\alpha} = \arctan\left(\frac{\alpha - \beta}{\alpha + \beta}\tan(kq)\right)\bigg|_{-\beta=\alpha} = \frac{\pi}{2} \qquad (13)$$

also consistent with (5). Then, in either case and consistent with (4), (6) would represent a scaled standing cosine wave for $\beta = \alpha$ and a scaled standing sine wave for $-\beta = \alpha$. Standing waves, while mathematically permitted, would have relativistic issues in addition to the aforementioned total reflection issue. Thus, the representation for the wave function (6) covers all solutions of physical interest of (2) propagating in the $+q$ direction with normalization $\alpha^2 - \beta^2 = 1$ (7).

If the neutrino and antineutrino are considered to form a Majorana pair of particles (an unsettled question), then the wave functions for the neutrino and antineutrino would be complex conjugates of each other. Under the Majorana hypothesis, the latent reflected wave $\beta\exp(-ikq)$ in (4) would be the wave function for a neutrino with amplitude β. In this case, (6) would represent the superposition of the wave functions of the Majorana neutrino and antineutrino upon each other. This is consistent with Pontecorvo's proposal [28] that a mixed particle consisting of part antineutrino and part neutrino may exist. Furthermore, the set of independent solutions $\{\overline{\psi}, \check{\overline{\psi}}\} = \{\exp(+ikq), \exp(-ikq)\} = \{\overline{\psi}, \psi\}$ that solve the SKGE, form a pair of Majorana solutions that are sufficient to solve the SKGE. Any solution, e.g. (4)–(6), of the SKGE formed from this pair by the superpositional principle would itself have a Majorana partner that would also be its complex conjugate. While the wave functions given by (4)–(6) are Pontecorvo "mixed" solutions [28], they are still specified herein as $\overline{\psi}$s of the $\overline{\nu}$ as determined by the directional characteristic $(+q)$ of the latent incident wave.

Let us briefly discuss how this ab initio calculation describes the evolution of the bispectral $\overline{\psi}_2$ during consummation of an IBD. The weak interaction is not a "force" *per se*. It does not cause an energy exchange among its participants. Rather, for purposes of this paper, it enables beta decay where a neutron decays into a proton, electron, and neutrino, which is the inverse of an IBD (1). Let us consider that the weak interaction occurs in a black box over the short range of the weak interaction between q_a, where the antineutrino initially encounters the weak interaction, and q_b where the antineutrino is absorbed by the target proton. The short range of the weak interaction is given by $q_b - q_a \approx 10^{-18}$ m, a value much smaller than the radius of the proton. Within the

black box $q_a < q < q_b$, the same set of independent solutions $\{\exp(+ikq), \exp(-ikq)\}$, which are sufficient to solve (2), are used to describe $\overline{\psi}_2$ while it is subject to the **forceless** weak interaction that precludes any energy exchange. In the absence of an energy exchange, the wave number k remains a constant in (4)–(6) during $\overline{\nu}_e$'s transit of the black box from q_a to q_b. But the coefficients $\{\alpha, \beta\}$ are changed! During the transit of $\overline{\nu}e$ from q_a to q_b in this ab initio calculation, the forceless weak interaction by W^+ exchange smoothly changes coefficients $\{\alpha, \beta\}|_{q_a} \to \{1, 0\}|_{q_b}$ while continuously maintaining the normalization $\alpha^2 - \beta^2 = 1$ of (7). In other words, the coefficients while inside the black box boundaries become variables $\{\alpha(q), \beta(q)\}_{q_a \leq q \leq q_b}$ that are explicitly still subject to the normalization

$$\alpha^2(q) - \beta^2(q) = 1, \quad q_a \leq q \leq q_b,$$

which is consistent with (7). A smooth transition of the coefficients from $\{\alpha(q_a), \beta(q_a)\}$ to $\{1, 0\}|_{q_b}$ with C^1 continuity would be sufficient to maintain C^1 continuity of the $\overline{\nu}_e$'s wave function as it evolves, during its transit of the black box with constant E and wave number k, from a bispectral $\overline{\psi}_2(q_a)$ to a unispectral $\exp(ikq_b)$ ready to be absorbed. At q_b, the output transmitted wave function of the black box will have become a unispectral wave function as given by

$$\overline{\psi}_2(q_a) = \alpha(q_a)\exp(ikq_a) + \beta(q_a)\exp(-ikq_a)$$
$$= [1 + \beta^2(q_a)]^{1/2}\exp(kq_a) + \beta(q_a)\exp(-ikq_a) \quad (14)$$
$$\xrightarrow{q \to q_b, \; \therefore \beta(q) \to 0} \exp(ikq_b), \quad q_a \leq q \leq q_b$$

under the influence of the exchange of the W^+ boson between the proton and antineutrino. In the extended black box, a provisional form for $\beta(q)$ with C^1 continuity during the transmutation of $\overline{\psi}$ from $\psi_2(q_a)$ to $\exp(ikq_b)$ in (14) is offered by

$$\beta(q) = \frac{\beta(q_a)}{2}\left[1 + \cos\left(\frac{q - q_a}{q_b - q_a}\pi\right)\right], \quad q_a \leq q \leq q_b.$$

Again, no energy is exchanged between the proton and antineutrino by the W^+ boson exchange. (If the transmitted wave function at q_b had not been unispectral $\exp(ikq)$, then its initial values at q_a would have been flavor incompatible $\overline{\nu}(q_a) \neq \overline{\nu}_e(q_a)$, which would have preempted an IBD. Consummated IBDs are rare events.) The transmitted unispectral wave function $\exp(ikq)$ is the wave function for $\overline{\nu}_e$ in (1). The normalization $\alpha^2 - \beta^2 = 1$ (7) specifies that the value of the amplitude of the transmitted unispectral wave function is 1, consistent with the assumptions for $\overline{\nu}_e$'s wave function for (1). The transmitted unispectral $\overline{\nu}_e$ is compatible with being absorbed by the proton consistent with (1). The function of the black box in the IBD process (to change the input bispectral wave function to an output unispectral wave function of amplitude 1 in a forceless manner for $\overline{\nu}_e$'s E never changes) has been completed with the $\overline{\nu}_e$ positioned at q_b, ready to be

absorbed with the target proton. The W^+ boson exchange has now been completed. The IBD carries on. The IBD completes consummation consistent with (1) where its parent particles, the proton and the unispectral antineutrino, are absorbed, and the IBD emits its daughter products, a positron and a neutron. The latent transmission coefficient T and reflective coefficient R of the black box for the weak interaction process are the expected

$$T = \frac{\alpha^2 - \beta^2}{\alpha^2} = \frac{1}{\alpha^2} \quad \text{and} \quad R = \frac{\beta^2}{\alpha^2}, \qquad (15)$$

where the coefficients $\{\alpha, \beta\}$ are their pre-weak interaction values.

Flavor compatibility for an IBD is determined by the boundary conditions $\{\overline{\psi}, \partial_q\overline{\psi}\}$ at the black box's input barrier interface q_a. The black box in this ab initio calculation renders a transmitted unispectral wave function $\exp(ikq)$, if and only if $\overline{\psi}_2$ has proper IBD initial values for the black box, $\{\overline{\psi}, \partial_q\overline{\psi}\}_{q=q_a}$.

Future research may refine the aforementioned description of the evolution of the antineutrino's wave function in the black box. If so, the principle of superposition of the wave functions of the latent incident and the latent reflected waves could still describe a generalized (14). For example, future research may find that the transmitted wave function of energy E from the black box should have coefficients $\{(1 + \beta_b^2)^{1/2}, \beta_b\}|_{q=q_b}$ with $\beta > 0$ for IBD absorption of the antineutrino. For a successful IBD, the black box model of the weak force would then transmute the incident wave function described by

$$[1 + \beta^2(q_a)]^{1/2}\exp(kq_a) + \beta(q_a)\exp(-ikq_a)$$
$$\xrightarrow{q \to q_b, \; \therefore \beta(q) \to \beta_b} \qquad (16)$$
$$(1 + \beta_b^2)^{1/2}\exp(ikq_b) + \beta_b\exp(-ikq_b)$$

where $q_a < q \leq q_b$. This generalizes (14) and would still describe a counterexample permitting massless neutrino oscillation. Eqs. (14) and (16) are analogous to the invariance of the Schwarzian derivative under a Möbius transformation in the quantum trajectory representation [14], [29].

Chirality and helicity are the same for massless leptons propagating with speed c. The quantum measure of helicity, normalized over a cycle of nonuniform propagation, for a massless antineutrino before encountering the black box, $q < q_a$, would by (4)–(6) be $\alpha^2 - \beta^2 = 1$, which is also the normalization (7). Upon completing the transit of the black box at q_b, the antineutrino, with $\overline{\psi} = \exp(ikq_b)$, would still have the helicity value of 1 conserving helicity (chirality). Thus, the interaction of the massless antineutrino with the black box would be reflectionless. This is consistent with (14) and (16). The concept of superimposing a latent reflected wave and the latent incident wave upon each other to achieve reflectionless transmission had initially been applied to an acoustical analogue [20].

The representation of $\overline{\psi}_2$ by (6) may be derived from the trigonometric form of (5) by using either Bohm's scheme for complex wave functions to render $\overline{\psi}_2$'s amplitude and phase [30] or by vector analysis. The amplitude $A_{\overline{\psi}} = [\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)]^{1/2}$ is recognized as a re-expressed law of cosines where the exterior angle argument $2kq$ is the supplement of $\pi - 2kq$ or

$$A_{\overline{\psi}} = \underbrace{[\alpha^2 + \beta^2 - 2\alpha\beta\cos(\pi - 2kq)]^{1/2}}_{\text{law of cosines}}$$
$$= \underbrace{[\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)]^{1/2}}_{\text{law of cosines for exterior angles}}.$$

For completeness, the phase is established [30] by $P_{\overline{\psi}}(q) = \arctan\{\Im[\overline{\psi}(q)]/\Re[\overline{\psi}(q)]\}$, which by (5) renders

$$P_{\overline{\psi}} = \arctan\left(\frac{\alpha - \beta}{\alpha + \beta}\tan(kq)\right). \tag{17}$$

Also for completeness, the phase is related to the quantum Hamilton's characteristic function (quantum reduced action) $\mathcal{W}$ by $P_{\overline{\psi}} = \mathcal{W}/\hbar$ [7], [10], [14]. The $\mathcal{W}$ has been shown to change values monotonically [14] implying that $P_{\overline{\psi}}$ also behaves monotonically.

The bispectral $\overline{\psi}_2$ as represented by (6) exhibits the superposition of the latent incident and reflected wave functions upon each other that are described by functions of q (4). The superposition induces a compound modulation in $\overline{\psi}_2$, which in turn induces nonuniform rectilinear propagation for massless neutrinos as shown in §3. PMNS theory achieves nonuniform rectilinear propagation in one dimension by superimposing three different masseigenstates within the neutrino [1]–[4]. Application of Eq. (6)-like representations have been made to study step barriers [18] and tunneling [19].

Before an IBD, $q \leq q_a$, the nonuniform propagation of the compoundly modulated $\overline{\psi}_2(q)$ with q can be examined more closely by considering the phase and amplitude modulations separately. The phase modulation may be described by the phase displacement between the phase of the bispectral $\overline{\psi}_2$ given by (6) and the phase kq of the corresponding unispectral wave function $\exp(ikq)$, which propagates rectilinearly with uniform motion. This phase displacement is a rotational displacement in complex $\overline{\psi}$-space between $\overline{\psi}_2(q)$ and the unispectral $\exp(ikq)$. The phase displacement due to phase modulation $Pm_{\overline{\psi}}$ may be expressed in units of radians as a function of phase kq, also in units of radians, as given by

$$Pm_{\overline{\psi}} = \arctan\left(\frac{\alpha - \beta}{\alpha + \beta}\tan(kq)\right) - kq, \quad q \leq q_a \tag{18}$$

where kq, which is also the phase of unispectral $\exp(ikq)$, is not restricted to its principal value.

The derivative of phase with respect to q, for the bispectral wave function (6) is given by [5]

$$\frac{\partial \arctan\left(\frac{\alpha - \beta}{\alpha + \beta}\tan(kq)\right)}{\partial q} = \frac{(\alpha^2 - \beta^2)k}{\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)}$$
$$= \frac{k}{\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)}. \tag{19}$$

Eq. (19) for the bispectral wave function exhibits nonuniform phase propagation that is periodic in q. The derivative of phase with respect to q remains positive definite for the denominator on the right side of (19) is always positive for all q by the Schwarzian inequality. Meanwhile, the corresponding derivative of phase for the unispectral wave function $\exp(ikq)$ is ik, which is constant and manifests uniform rectilinear propagation. For completeness in the quantum trajectory representation, the derivative of phase with regard to q renders the conjugate momentum $\partial_q \mathcal{W}$ divided by $\hbar$ [8]–[14].

The relative amplitude dilation $Am_{\overline{\psi}}$ due to amplitude modulation $A_{\overline{\psi}}$ of (6) or (8), relative to $(\alpha^2 + \beta^2)^{1/2}$, is defined to be a dimensionless variable that is a function of phase kq and given by

$$Am_{\overline{\psi}} \equiv \frac{[\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)]^{1/2} - (\alpha^2 + \beta^2)^{1/2}}{(\alpha^2 + \beta^2)^{1/2}}$$
$$= \left[1 + \frac{2\alpha\beta\cos(2kq)}{\alpha^2 + \beta^2}\right]^{1/2} - 1, \quad q \leq q_a. \tag{20}$$

Any finite $\beta = (\alpha^2 - 1)^{1/2}$ is sufficient to cause $\overline{\psi}_2$ to generate nonuniform rectilinear motion consistent with the compound modulation implied by (18) and (20).

As the wave function $\overline{\psi}_2$ for the antineutrino must be C^1 continuous until absorbed in an IBD, the behavior of its derivative $\partial_q \overline{\psi}_2$ must also be considered. If the dividend of $\partial_q \overline{\psi}_2 / \overline{\psi}_2$ were a constant or independent of q, then neutrino oscillation would not be supported as previously noted. From (4)–(6), the derivative of the bispectral wave function $\partial_q \overline{\psi}_2$ is given by

$$\partial_q \overline{\psi}_2 = ik[\alpha\exp(ikq) - \beta\exp(-ikq)]$$
$$= k[(\alpha - \beta)\cos(kq) - i(\alpha + \beta)\sin(kq)]\exp(i\pi/2)$$
$$= k\underbrace{[\alpha^2 + \beta^2 - 2\alpha\beta\cos(2kq)]^{1/2}}_{\text{law of cosines}} \tag{21}$$
$$\times \exp\left[i\arctan\left(\frac{\alpha + \beta}{\alpha - \beta}\tan(kq)\right) + i\frac{\pi}{2}\right].$$

A difference between (4)–(6) for $\overline{\psi}_2$ and (21) for $\partial_q \overline{\psi}_2$ is the change of the sign of β and the phase shift $\pi/2$. A finite β by (4) and (21) ensures that

$$\frac{\partial_q \overline{\psi}_2(q)}{\overline{\psi}_2(q)} = ik\left(\frac{\alpha\exp(ikq) - \beta\exp(-ikq)}{\alpha\exp(ikq) + \beta\exp(-ikq)}\right) \tag{22}$$

would be a variable of q in contrast to the unispectral case (3). The bispectral $\overline{\psi}_2(kq)$ propagates in a nonuniform manner

that facilitates neutrino oscillation without the need for mass-eigenstates of PMNS theory.

There is an alternative expression for $\partial_q \overline{\psi}_2(kq)$ that conveniently shows its relation to $\overline{\psi}(kq - \pi/2)$. This relation is shown by (4) and (21) to be

$$\begin{aligned} \partial_q \overline{\psi}_2(kq) &= ik[\alpha \exp(ikq) - \beta \exp(-ikq)] \\ &= k\{\alpha \exp[i(kq + \pi/2)] + \beta \exp[-i(kq + \pi/2)]\} \quad (23) \\ &= k\,\overline{\psi}_2(kq + \pi/2). \end{aligned}$$

Eq. (23) can be generalized to

$$\begin{aligned} \partial_q \overline{\psi}_2(kq) &= k\overline{\psi}_2(kq + n_1\pi), \\ n_1 &= \pm 1/2, \pm 3/2, \pm 5/2, \cdots. \end{aligned} \quad (24)$$

where n_1 is bound by the antineutrino's creation point and the point q_a where an IBD commences. The bispectral derivative $\partial_q \overline{\psi}_2$ by (21)–(24), like $\partial_q \overline{\psi}_1$, is also a solution of the SKGE.

The derivative of the bispectral wave function is compoundly modulated. Its amplitude $A_{\overline{\psi}'}$ and phase $P_{\overline{\psi}'}$ are respectively given by

$$A_{\overline{\psi}'} = k[\alpha^2 + \beta^2 - 2\alpha\beta \cos(2kq)]^{1/2}, \quad q \le q_a \quad (25)$$

and

$$P_{\overline{\psi}'} = \arctan\left(\frac{\alpha + \beta}{\alpha - \beta} \tan(kq)\right) + \frac{\pi}{2}, \quad q \le q_a. \quad (26)$$

Its relative amplitude dilation $Am_{\overline{\psi}'}$ due to amplitude modulation and its phase displacement (a rotation) due to phase modulation $Pm_{\overline{\psi}'}$ for $\partial_q \overline{\psi}_2(kq)$ are given respectively by

$$Am_{\overline{\psi}'} = k\left[1 - \frac{2\alpha\beta \cos(2kq)}{\alpha^2 + \beta^2}\right]^{1/2} - k, \quad q \le q_a \quad (27)$$

and

$$Pm_{\overline{\psi}'} = \arctan\left(\frac{\alpha + \beta}{\alpha - \beta} \tan(kq)\right) - kq, \quad q \le q_a. \quad (28)$$

The dilations and rotations of (27) and (28) for $\partial_q \overline{\psi}_2(kq)$ are analogous to those for $\overline{\psi}_2$, (20) and (18) respectively. While $\partial_q \overline{\psi}_2(kq)$ has compound modulation with the same period (oscillation cycle) as that of the associated $\overline{\psi}_2(kq)$, the of dilations and rotations differ by being out of phase, *cf.* (6) and (21)–(28). The relative amplitude dilation and phase rotation of $\partial_q \overline{\psi}_2(kq)$ are opposite to those of $\overline{\psi}_2(kq)$. This is desirable for flavor oscillation.

Let us now examine the measurement of momentum p for the bispectral antineutrino. The applicable quantum momentum operator herein is $\frac{\hbar}{i}\partial_q$. The orthodox measurement of momentum of the bispectral $\overline{\psi}_2$ with box normalization is

over one repetitive cycle. This box length is π/k. The momentum of $\overline{\psi}_2$, using (4), (7) and (21), is given by

$$\begin{aligned} p &= \frac{\int_0^{\pi/k} \overline{\psi}_2^\dagger(q) \frac{\hbar}{i}\partial_q \overline{\psi}_2(q)\,dq}{\int_0^{\pi/k} \overline{\psi}_2^\dagger(q)\overline{\psi}_2(q)\,dq} \\ &= \hbar \frac{k \int_0^{\pi/k} [\alpha^2 - \beta^2 + 2\alpha\beta \sin(2kq)]\,dq}{\int_0^{\pi/k} [\alpha^2 + \beta^2 + 2\alpha\beta \cos(2kq)]\,dq} \quad (29) \\ &= \hbar \frac{(\alpha^2 - \beta^2)\pi}{(\alpha^2 + \beta^2)\pi/k} = \frac{\hbar k}{\alpha^2 + \beta^2}. \end{aligned}$$

An orthodox measurement of momentum of the bispectral antineutrino (29) is a constant and positive definite, *i.e.* $p > 0$, in the direction of latent incident wave (4). This is consistent with the quantum trajectory representation where the quantum reduced action $\mathcal{W}$ changes monotonically [14].

Let us extend our examination of p to find under what conditions $[\alpha^2 - \beta^2 + 2\alpha\beta \sin(2kq)]$, the integrand in the numerator in (29), becomes negative over any portions of its repetitive cycle. The particular point of interest for investigation is $q = 3\pi/(4k)$ where the integrand becomes

$$[\alpha^2 - \beta^2 + 2\alpha\beta \sin(2kq)]_{q=3\pi/(4k)} = \overbrace{\alpha^2 - \beta^2}^{=1} - 2\alpha\beta. \quad (30)$$

For $|\beta|$ sufficiently small, (30) would be positive; sufficiently large, negative. The $|\beta|$ for which (30) is nil marks the upper bound where $[\alpha^2 - \beta^2 + 2\alpha\beta \sin(2kq)]$, the integrand, is never negative. Because $-\beta^2$ is a negative quantity, the Schwarz inequality is not applicable to (30). The right side of (30) becomes nil for

$$2\alpha\beta = 1. \quad (31)$$

The particular values of α and $|\beta|$ that satisfy both Eqs. (7) and (31) are identified by $\alpha_{\text{threshold}}$ and $|\beta_{\text{threshold}}|$. The threshold coefficients separate $\alpha, |\beta|$-space into two domains: one where the integrand is always positive-definite; the other, not always positive consistent with the value of $\sin(2kq)$ in (29). Eq. (7) for normalization, $\alpha^2 - \beta^2 = 1$, and (31) are sufficient to resolve $\alpha_{\text{threshold}}$ and $|\beta_{\text{threshold}}|$ by algebraic means. The solutions for the threshold coefficients are

$$\{\alpha_{\text{threshold}}, \beta_{\text{threshold}}\} = \left\{\left(\frac{2^{1/2}+1}{2}\right)^{1/2}, \left(\frac{2^{1/2}-1}{2}\right)^{1/2}\right\}. \quad (32)$$

The logic relationship

$$\alpha < / > \alpha_{\text{threshold}} \iff |\beta| < / > |\beta_{\text{threshold}}|$$

between α and β follows. If $|\beta| < |\beta_{\text{threshold}}|$, then the integrand $\overline{\psi}_2^\dagger(q)(\hbar/i)\partial_q \overline{\psi}_2(q)$ of (29) would always be positive (in the direction of the latent incident wave of (4)) for all q throughout the repetitive oscillation cycle. If $|\beta| > |\beta_{\text{threshold}}|$, then for some q, but not a preponderance of q of the repetitive oscillation cycle, the integrand $\overline{\psi}^\dagger (\hbar/i)\partial_q \overline{\psi}_2$ would be negative (in
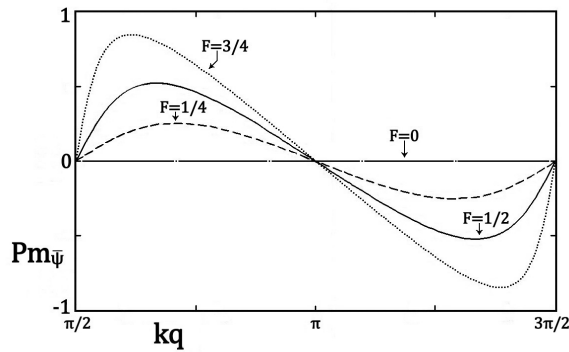
Fig. 1: The phase displacement due to phase modulation $Pm_{\overline{\psi}}$ as a function of kq over a Riemann sheet for selected values of F. Both $Pm_{\overline{\psi}}$ and kq are exhibited in units of radians.

the direction of the latent reflected wave of (4)). Nevertheless, even if $|\beta| > |\beta_{\text{threshold}}|$, the orthodox measure for momentum would still remain valid, for (29) yields positive momentum as $\alpha^2 - \beta^2 = 1 > 0$.

3 Examples

Let us now illustrate with didactic examples how a bispectral wave function facilitates massless flavor oscillation. We consider the contributions of phase and amplitude modulations separately. These contributions are examined for the selected cases given by

$$(\alpha, \beta) = (1, 0), (4/15^{1/2}, 1/15^{1/2}), (2/3^{1/2}, 1/3^{1/2}),$$
$$(4/7^{1/2}, 3/7^{1/2}). \tag{33}$$

These cases are compliant with normalization $\alpha^2 - \beta^2 = 1$ (7). The selected cases may be identified for convenience by the fraction $F \equiv \beta/\alpha = (\alpha^2 - 1)^{1/2}/\alpha = \beta/(1 - \beta^2)^{1/2}$. Also, F is related to the reflection coefficient (15) for $F = R^{1/2}$. The fractions F for the selected cases with respect to (33) are given by

$$F = 0, 1/4, 1/2, 3/4. \tag{34}$$

Comparisons of the effects of either phase or amplitude modulations among the selected cases of F are developed as a function of phase kq measured in radians.

The value $F = 0$ represents a unispectral wave function, which precludes massless flavor oscillation. The unispectral $F = 0$ is still included for comparison to the bispectral Fs where $F = 1/4, 1/2, 3/4$. For comparison, the value $F_{\text{threshold}}$ for $2\alpha\beta = 1$ with normalization $\alpha^2 - \beta^2 = 1$, which establishes F's upper bound for no reversals of sign of the integrand $\overline{\psi}_2^{\dagger}(\hbar/i)\partial_q \overline{\psi}_2$ as a function of q (32) is given by

$$F_{\text{threshold}} = \frac{\beta_{\text{threshold}}}{\alpha_{\text{threshold}}} = \left(\frac{2^{1/2} - 1}{2^{1/2} + 1}\right)^{1/2} = 2^{1/2} - 1$$
$$= \frac{1}{2^{1/2} + 1} = 0.41421356\cdots.$$

We first consider phase modulation. The phase displacements $Pm_{\overline{\psi}}$ of (18) as a function of kq, where kq is also the phase of $\overline{\psi}$, are exhibited for the various values of F on Fig. 1 over the extended Riemann sheet $\pi/2 \leq kq \leq 3\pi/2$ of the arc tangent function on the right side of (6). The phase duration of the Riemann sheet is consistent with box normalization of $\overline{\psi}_2$. Each extended Riemann sheet specifies an oscillation cycle. Fig. 1 exhibits one cycle for phase modulation $Pm_{\overline{\psi}}$ over a Riemann sheet. The cycle of $Pm_{\overline{\psi}}$ for bispectral Fs has one concave segment and one convex segment. The cycle is repetitive over other Riemann sheets. As expected, a $Pm_{\overline{\psi}}$ for the unispectral F renders the horizontal straight line $Pm_{\overline{\psi}} = 0$. Thus, the unispectral case prohibits phase modulation, which does not facilitate flavor oscillation. The absolute value of $Pm_{\overline{\psi}}$ for $kq \neq \pi/2, \pi, 3\pi/2$ is shown on Fig. 1 to increase with increasing F. At $kq = \pi/2, \pi, 3\pi/2$, the phase difference $Pm_{\overline{\psi}} = 0$ for all F. These points $kq = \pi/2, \pi, 3\pi/2$ for $F \neq 0$, are inflection points of $Pm_{\overline{\psi}}$ with nil curvature, which are between $Pm_{\overline{\psi}}$'s alternating concave and convex segments. At these inflection points, $|Pm_{\overline{\psi}}(q)|$ attains its maximum slope (rate of change with kq). Had Fig. 1 included the standing-wave case where $F = 1$, then, consistent with (10) and (11), it would have generated a straight line from $Pm_{\overline{\psi}}(kq) = (\pi/2, \pi/2)$ to $(-\pi/2, 3/\pi/2)$ on an extended Fig. 1. Had the cases $F = -1/4, -1/2, -3/4$ been examined instead (e.g. the values of F for the analogous phase differences for $\partial_q \overline{\psi}_2$ would be negative), then Fig. 1 would have changed its exhibition of the antisymmetric phase modulation from the first-and-third (upper/left-and-lower/right) quadrants to the second-and-fourth of Fig. 1. The phase modulation $Pm_{\overline{\psi}}$ is antisymmetric within the Riemann sheet for

$$Pm_{\overline{\psi}}(\pi - kq) = -Pm_{\overline{\psi}}(\pi + kq), \quad 0 < q < \pi/2.$$

Each extended Riemann sheet contains one cycle of $Pm_{\overline{\psi}}$ for the bispectral $\overline{\psi}_2$.

For the amplitude modulation, $Am_{\overline{\psi}}$ is examined for $F = 0, 1/4, 1/2, 3/4$. Again, $F = 0$ represents the unispectral case, which does not support flavor oscillation. The amplitude modulations are exhibited on Fig. 2. Positive differences on Fig. 2 represent a dilation that is an expansion; negative differences, a contraction. The absolute values of $Am_{\overline{\psi}}$ for $kq \neq 3\pi/4, \pi/4$ are shown on Fig. 2 to increase with increasing F. In Fig. 2, $Am_{\overline{\psi}}$ for bispectral F is symmetric with its convex segments disjointed on the Riemann sheet. In comparing Figs. 1 and 2 for bispectral $F = 1/4, 1/2, 3/4$, either the
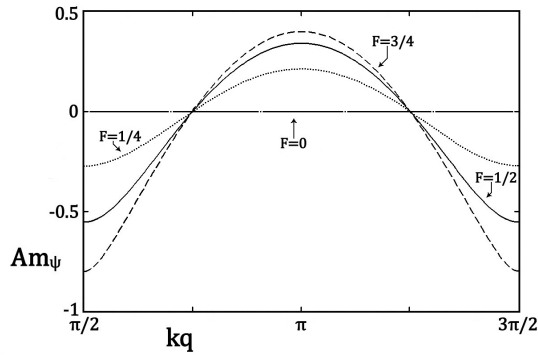
Fig. 2: The relative amplitude dilation $Am_{\overline{\psi}}$ as a function of kq over a Riemann sheet for selected values of F. $Am_{\overline{\psi}}$ is dimensionless, and kq is exhibited in units of radians.
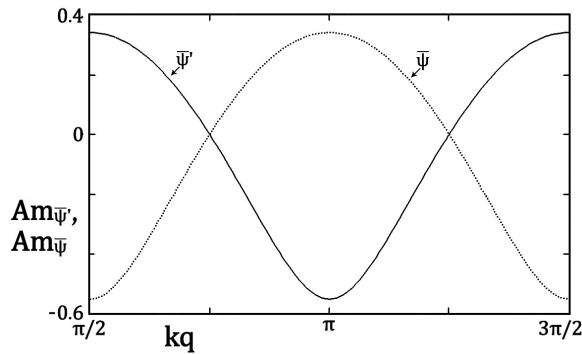

Fig. 3: The relative amplitude dilations due $Am_{\overline{\psi}}$ and $Am_{\overline{\psi}'}$ as functions of kq over a Riemann sheet for $F = 1/2$. For an unbiased $Am_{\overline{\psi}'}$, $k = 1$ to facilitate comparison to dimensionless $Am_{\overline{\psi}}$. The amplitude modulations are dimensionless, and kq is exhibited in units of radians.

$Pm_{\overline{\psi}}$ or the $Am_{\overline{\psi}}$ has an extremum where the other is nil. This ensures that at least one type of modulation of $\overline{\psi}_2$ is changing for all q on the extended Riemann sheet $\pi/2 \le kq \le 3\pi/2$. A local maximum rate of change of a modulation occurs at its zero-crossings where the modulation has inflection points between concave and convex segments as shown by Figs. 1 and 2. The greater (lesser) rate of change of modulation implies the greater (lesser) opportunity for flavor oscillation. The modulation extrema, where the rate of change of a particular modulation is nil, are isolated phase (kq) points where that particular modulation does not contribute to neutrino oscillation.

A comparison between the amplitude modulation $Am_{\overline{\psi}}$ of the bispectral $\overline{\psi}_2$ (6) and the amplitude modulation $Am_{\overline{\psi}'}$ of the associated bispectral $\partial_q \overline{\psi}_2$ (21) are presented in Fig. 3 for the particular values $F = 1/2$, and $k = 1$. As $Am_{\overline{\psi}'}$ by (25) has a linear factor k while $Am_{\overline{\psi}}$ does not, the choice $k = 1$ makes

Fig. 3 unbiased. The amplitude modulations $Am_{\overline{\psi}}$ and $Am_{\overline{\psi}'}$ exhibit the same repetitive periodicity but are displaced in phase (kq) by the constant $\pi/2$ radians. This kq displacement increases the opportunity for neutrino oscillation for $Am_{\overline{\psi}}(kq)$ is positive (negative) where $Am_{\overline{\psi}'}(kq)$ is negative (positive).

The ratio of amplitudes of $\partial_q \overline{\psi}_2(kq)$ relative to $\overline{\psi}_2(kq)$ by (6) and (21) is given as a function of phase (kq) in fractional form by

$$
\left| \partial_q \overline{\psi}_2(kq) \right| : \left| \overline{\psi}_2(kq) \right| \rightsquigarrow \underbrace{\frac{\left| \partial_q \overline{\psi}_2(kq) \right|}{\left| \overline{\psi}_2(kq) \right|}}_{\text{fractional form}} = \frac{A_{\overline{\psi}'}(kq)}{A_{\overline{\psi}}(kq)})
$$
$$
= k \left(\frac{\alpha^2 + \beta^2 - 2\alpha\beta \cos(2kq)}{\alpha^2 + \beta^2 + 2\alpha\beta \cos(2kq)} \right)^{1/2} . \tag{35}
$$

On the extended Riemann sheet $\pi/2 \le kq \le 3\pi/2$, the ratio $A_{\overline{\psi}'}(kq) : A_{\overline{\psi}}(kq)$ for $F = 1/2$ by (33)–(35) has maxima of $3k$ at $kq = \pi/2$, $3\pi/2$; has a minimum of $k/3$ at $kq = \pi$; and equals k at $kq = 3\pi/4$, $5\pi/4$ in accordance with (35). The values of the extrema of ratio in fractional form (35) may be generalized and are given on this extended Riemann sheet by

$$
\frac{A_{\overline{\psi}'}(kq)}{A_{\overline{\psi}}(kq)} \bigg|_{\text{maximum}} = k \frac{\alpha + \beta}{\alpha - \beta} \text{ at } kq = \frac{\pi}{2}, \frac{3\pi}{2}
$$

and

$$
\frac{A_{\overline{\psi}'}(kq)}{A_{\overline{\psi}}(kq)} \bigg|_{\text{minimum}} = k \frac{\alpha - \beta}{\alpha + \beta} \text{ at } kq = \pi.
$$

The nature of (35) implies that its logarithmic presentation would exhibit for unbiased $k = 1$ a periodic antisymmetry within the extended Riemann sheet $\{\pi/2 \le kq \le 3\pi/2\}$ given by

$$
\ln \left(\frac{A_{\overline{\psi}'}(kq)}{A_{\overline{\psi}}(kq)} \right) = -\ln \left(\frac{A_{\overline{\psi}'}(kq \pm \pi/2)}{A_{\overline{\psi}}(kq \pm \pi/2)} \right), \text{ for } k = 1.
$$

The variation of the ratio (35) is one of the factors that facilitate flavor oscillation. On the other hand, the corresponding ratio for the unispectral case ($F = 0$) is the constant k for all q.

A comparison of (9) and (26) shows the relationship between $P_{\overline{\psi}}(kq)$ and $P_{\overline{\psi}'}(kq)$ is that the sign of β has changed (also the sign of the associated F would change). Therefore $P_{\overline{\psi}'}(kq) - \pi/2$ and $P_{\overline{\psi}}(kq)$ are a half-cycle out of phase. While the undulations of $P_{\overline{\psi}'}$ and $P_{\overline{\psi}}$ when summed are in opposition, their difference is reinforced. Their changing difference is another factor enabling flavor oscillation. The relative phase difference $\triangle P_{\overline{\psi}', \overline{\psi}}(kq)$ in radians between $Pm_{\overline{\psi}'}$ and $Pm_{\overline{\psi}}$ is reinforced for they are out of phase as shown by

$$
\triangle P_{\overline{\psi}', \overline{\psi}}(kq) = P_{\overline{\psi}'}(kq) - P_{\overline{\psi}}(kq)
$$
$$
= P_{\overline{\psi}}(kq + \pi/2) + \pi/2 - P_{\overline{\psi}}(kq). \tag{36}
$$
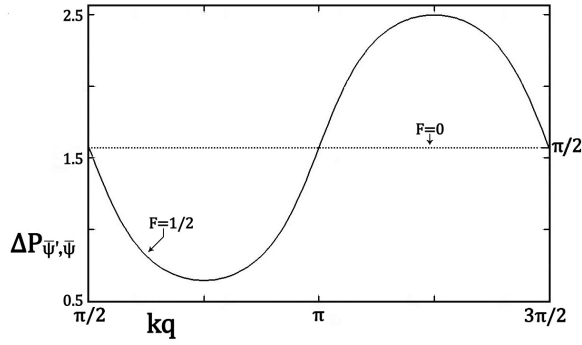
Fig. 4: The Phase difference $\triangle P_{\overline{\psi}',\overline{\psi}}(kq)$ as a function of kq over a Riemann sheet for $F = 0$, $1/2$. Both $\triangle P_{\overline{\psi}',\overline{\psi}}(kq)$ and kq are exhibited in units of radians.

The relative phase difference $\triangle P_{\overline{\psi}',\overline{\psi}}(kq)$ is exhibited on Fig. 4 for $F = 1/2$ and $F = 0$ (the unispectral case). For the bispectral case, Fig. 4 also exhibits coherent reinforcement of the undulations of $P_{\overline{\psi}'}$ and $P_{\overline{\psi}}$ of $\triangle P_{\overline{\psi}',\overline{\psi}}(kq)$ consistent with (23). Larger undulations increase the opportunity for flavor oscillations.

The two factors, the ratio of amplitudes and the phase difference, describe the relative relationship between $\partial_q \overline{\psi}$ and $\overline{\psi}$ as a function of phase kq. The ratio of amplitudes (35) and the phase difference of Fig. 4 each complete one cycle on an extended Riemann sheet, $e.g.$ $\pi/2 < kq < 3\pi/2$. However, their respective extrema are displaced by a quarter cycle $\pi/4$ from each other. The phase difference $\triangle P_{\overline{\psi}',\overline{\psi}}(kq)$ has extrema on the extended Riemann sheet at $kq = 3\pi/4$, $5\pi/4$ while the ratio $A_{\overline{\psi}'} : A_{\overline{\psi}}(kq)$ has extrema at $kq = \pi/2$, π, $3\pi/2$. Where one factor has an extremum at some particular kq, the other factor has an inflection point there. And where one factor has an inflection point, the other has an extremum. A local extremum for a factor implies that the factor has a local nil in facilitating flavor oscillation while the other factor having an inflection point implies a local peak in facilitating flavor oscillation. Furthermore, where one factor's support for flavor oscillation decreases, the other factor's support increases. Thus, the two factors complement each other to ensure that the bispectral antineutrino can facilitate possible flavor oscillation for some interaction throughout its repetitive cycle.

Both phase and amplitude modulations exhibit the same kq periodicity on Figs. 1–4. This may be shown by trigonometry for the general situation. Periodicity of phase modulation (19) is consistent with the extended Riemann sheet of the arc tangent,

$$(2n - 1)\pi/2 \leq kq \leq (2n + 1)\pi/2, \; n = 0, \pm 1, \pm 2, \cdots .$$

Hence, $Pm_{\overline{\psi}}(kq) = Pm_{\overline{\psi}}(kq + \pi)$. Periodicity of amplitude modulation (20) is consistent with the argument $2kq$ of the cosine term in the law of cosines completing its cycle 2π.

Periodicity of $Am_{\overline{\psi}}$ is also given by

$$Am_{\overline{\psi}}(kq) = Am_{\overline{\psi}}(kq + n\pi), \; n = \pm 1, \pm 2, \pm 3, \cdots .$$

For completeness, the quantum trajectory representation also has the same kq periodicity [5].

4　Discussion

Compound modulation makes $\partial_q \overline{\psi}_2 / \overline{\psi}_2$ a periodic variable in phase kq and spatially periodic for a given k. The phase and amplitude modulations complement each other for they are a quarter-cycle out of phase with each other as shown by Figs. 1 and 2. The modulations of $\overline{\psi}_2$ and $\partial_q \overline{\psi}_2$ supplement each other. The amplitude modulation induces continuous dilations with respect to phase kq of the $\partial_q \overline{\psi}_2(q)$ and $\overline{\psi}_2(q)$ differently by (25) and (8) respectively. The dilations of $\partial_q \overline{\psi}_2(q)$ and $\overline{\psi}_2(q)$ are opposed: where one is an expansion; the other is a contraction. These amplitude modulations being in opposition increase the amount of dilation (either expansion or contraction) of the ratio $\left| \partial_q \overline{\psi}_2(kq) \right| : \left| \overline{\psi}_2(kq) \right|$ with respect to phase kq as exhibited by (35) and Fig. 3. This increases the opportunity for neutrino oscillation. Meanwhile, phase modulation induces continuous rotations with respect to phase kq of $Pm_{\overline{\psi}}(q)$ (18) and $Pm_{\overline{\psi}'}(q)$ (28). These rotational displacements are opposed: where one rotation is clockwise; the other, counterclockwise. This opposition in rotations enlarges $\triangle P_{\overline{\psi}',\overline{\psi}}(kq)$ as exhibited by (36) and Fig. 4. This opposition between the behavior of $\overline{\psi}_2(q)$ and its derivative is typical of well behaved functions undergoing periodic motion. Note that either phase or amplitude modulation, by itself, could facilitate neutrino oscillation of the bispectral antineutrino. Together, they increase the opportunity for oscillation.

The transmutation of coefficients $\{\alpha, \beta\} \rightarrow \{1, 0\}$ of (14) by the weak interaction nulls out the compound modulation of $\overline{\nu}_e$'s wave function without any exchange of energy. This is shown for phase modulation on Fig. 1 and for amplitude modulation on Fig. 2 where modulation effects decrease with decreasing absolute values of $|F|$ and are completely nulled at $|F| = 0$.

The periodic, nonuniform propagation by a massless antineutrino results in flavor oscillations where the antineutrino in a particular phase (kq) segment within an oscillation cycle may execute a flavor-compatible current interaction with C^1 continuity of its wave function. Future work may show that these segments for various flavors $\{\overline{\nu}_e, \overline{\nu}_\mu, \overline{\nu}_\tau\}$ may be disjointed, and the segments for the flavors may not densely fill the oscillation cycle.

Should the segments for the active flavors $\{\overline{\nu}_e, \overline{\nu}_\mu, \overline{\nu}_\tau\}$ not densely fill the oscillation cycle, then the voids of the oscillation cycle would be locations where the antineutrino is inactive and would behave as the elusive sterile antineutrino $\overline{\nu}_s$ [31], [32]. By precept, the sterile antineutrino was hypothesized to be subject only to gravity and explicitly not to

the weak interaction. The MiniBooNE Collaboration has recently inferred its existence from experiment [31], but such existence has not yet been independently confirmed by other ongoing experiments [32]. As the hypothetical sterile antineutrino would not partake in charged current interactions, the voids in the oscillation cycle could manifest the existence of this hypothetical sterile antineutrino. This hypothetical sterile antineutrino, by (2)–(6), could be massless and have a bispectral wave function. As this hypothetical bispectral sterile antineutrino could propagate nonuniformly, it would oscillate in flavor to become an active antineutrino $\{\overline{\nu}_e, \overline{\nu}_\mu, \overline{\nu}_\tau\}$. Flavor oscillation of the sterile antineutrino would imply that it would have the same right handedness of the active antineutrinos. Again, this support for the existence of the sterile antineutrino is predicated on the existence of voids in the oscillation cycle.

The orthodox measurement of the momentum operator $\frac{\hbar}{i}\partial_q$ acting on a bispectral antineutrino over a box length, which is consistent with an oscillation cycle, has been shown by (29) to give a finite positive momentum in the direction of the latent incident wave (4). An IBD event is a good way to observe antineutrinos for the antineutrino reacts only to gravity and the weak interaction. Observed momentum, in principle, need not be averaged over a box length. Should future work find that box normalization is too coarse, then restricting the absolute value of β to $|\beta| \leq |\beta_{\text{threshold}}| = [(2^{1/2}-1)/2]^{1/2}$ (32) would maintain positive momentum for the bispectral antineutrino throughout the oscillation cycle, *i.e.*

$$\overline{\psi}_2^\dagger(q)\frac{\hbar}{i}\partial_q\overline{\psi}_2(q) > 0$$

by (30)–(32) for all q within the box normalization.

Future work may also show that the different charged or neutral current interactions may scramble the flavors. In other words, the antineutrino flavors may be interaction dependent where the values of $\overline{\psi}$ and $\partial_q\overline{\psi}$ for some given E at a point q_0 may specify an antineutrino of a particular flavor for an interaction while concurrently at q_0 also specifying a different flavor associated with another different interaction. This would cause the segments for the various flavors of the oscillation cycle to overlap.

Future work may also yield a better understanding of IBDs and the weak force. Nevertheless, the concept of a bispectral wave function representation should be robust enough to adjust assumptions and still facilitate flavor oscillation by a massless antineutrino.

5 Findings and conclusions

The principal finding is the existence of a wave function representation for massless neutrino oscillation of flavor, which is a counterexample to PMNS theory's finding that $m > 0$. The wave function representation for $m = 0$ is compatible with an orthodox interpretation of the bispectral wave function, $\overline{\psi}_2$. One spectral component represents the embedded

latent incident wave function for an IBD; the other, the embedded latent reflected wave function. Such a bispectral wave function is capable of flavor oscillations without any need for mass-eigenstates, which confirms that PMNS theory is not the exclusive theory for neutrino oscillation. Once created, a bispectral, massless antineutrino, with super-threshold energy ($E > 1.806$ MeV), has the possibility by flavor oscillation to initiate an IBD.

The co-principal finding, which is extra to the massless oscillation finding, is that the forceless weak interaction for this oscillation model transmutes the wave function of the antineutrino from bispectral to unispectral. There is no energy exchange during the transmutation for the weak interaction is forceless. In general, the weak interaction can transmute the wave function to a different superposition of its set of independent solutions without any exchange of energy.

The first secondary finding is that flavor oscillations are compatible with classifying neutrinos to be Majorana leptons.

The second secondary finding is that the elusive sterile neutrino may be just where the antineutrino is in a location, q, in the oscillation cycle where its values $\{\overline{\psi}_2, \partial_q\overline{\psi}_2\}|_q$ are incompatible initial values for initiating a current interaction of any flavor there (sterile is not a flavor). This finding is predicated upon the existence of such a location in the oscillation cycle.

The third secondary finding establishes a relationship between the amplitude β of the latent embedded reflected wave and the opportunity to observe negative momentum, *i.e.*, $\overline{\psi}_2^\dagger(q)\frac{\hbar}{i}\partial_q\overline{\psi}_2(q) < 0$. There exists a $\beta_{\text{threshold}}$ for which, if $|\beta| < |\beta_{\text{threshold}}|$, then $\overline{\psi}_2^\dagger(q)\frac{\hbar}{i}\partial_q\overline{\psi}_2(q) > 0$ for all q before an IBD. For cases of super-threshold $|\beta|$, the orthodox quantum measurement of momentum over one repetitive box length would still yield positive momentum (29).

The fourth secondary finding confirms the similar prediction for massless neutrino oscillation by the less familiar quantum trajectory representation of quantum mechanics [5]. This finding also substantiates that wave mechanics and quantum trajectories are equivalent for free particles [7], [33]. In addition, incisive insights rendered by the wave function representation complement those of the trajectory representation to substantiate massless neutrino oscillation.

A tertiary finding supports Pontecorvo's suggestion [28] that a neutrino may be composed of a mixture of neutrino and antineutrino components.

In conclusion, massless neutrino oscillation implies the validity of the standard model to consider neutrinos to be massless.

A co-conclusion is that the forceless weak interaction prepares the antineutrino for interaction with other particles by transmuting the antineutrino's wave function. The transmutation changes the wave function in this ab initio calculation from a bispectral wave function to a unispectral wave function $\exp(ikq)$ without an exchange of energy. Conversely, the

wave function of the antineutrino manifests the effects of the forceless weak interaction by a change in the superposition of its independent solutions for a given energy.

A secondary conclusion is the confirmation of the similar prediction of the validity of the standard model by the quantum trajectory representation, which substantiates that such a prediction is not an anomaly of the quantum trajectory representation.

Acknowledgement

References

1. Pontecorvo B. Mesonium and antimesonium. *Sov. Phys. JETP*, 1958, v. 6, 429–31; in Russian: *Zh. Eksp. Teor. Fiz.*, 1957, v. 33, 549–57.

2. Pontecorvo B. Neutrino experiment and the problem of electronic charge. *Sov. Phys. JETP*, 1968, v. 26, 984–8; in Russian: *Zh. Eksp. Teor. Fiz.*, 1967, v. 53, 1717–1725.

3. Maki B., Nakagawa N. and Sakata S. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 1962, v. 28, 870–80.

4. Mohapatra R. N. and Smirnov A. Y. Neutrino mass and new physics. *Ann. Rev. Nucl. Part. Sci.*, 2006, v. 56, 569–628. arXiv: hep-ph/0603118.

5. Floyd E. R. Neutrino oscillations with nil mass. *Found. Phys.*, 2017, v. 47, 42–60. arXiv: 1607.05577.

6. Hecht C. E. and Mayer J. E. Extension of the WKB equation. *Phys. Rev.*, 1953, v. 106, 1156–60.

7. Messiah A. Quantum Mechanics, v. I, North Holland, Amsterdam, 1961, pp. 222–8.

8. Floyd E. R. Modified potential and Bohm's quantum potential. *Phys. Rev. D*, 1982, v. 26, 1339–47.

9. Floyd E. R. Arbitrary initial conditions of hidden variables. *Phys. Rev. D*, 1984, v. 29, 1842–4.

10. Floyd E. R. Closed form solutions for the modified potential. *Phys. Rev. D*, 1986, v. 34, 3246–9.

11. Floyd E. R. Classical limit of the trajectory representation of quantum mechanics, loss of information and residual indeterminacy. *Int. J. Mod. Phys. A*, 1999, v. 14, 1111–24. arXiv: quant-ph/9708026.

12. Faraggi A. E. and Matone M. Quantum mechanics from an equivalence principal. *Phys. Lett. B*, 1999, v. 450, 34–40. arXiv: hep-th/9705108.

13. Carroll R. Some remarks on time, uncertainty and spin. *J. Can. Phys.*, 1999, v. 77, 319–25. arXiv: quant-ph/9903081.

14. Faraggi A. E. and Matone M. The equivalence postulate of quantum mechanics. *Int. J. Mod. Phys. A*, 2000, v. 15, 1869–2017. arXiv: hep-th/9809127.

15. Bertoldi G., Faraggi A. E. and Matone M. Equivalence principal, higher dimensional Möbius Group and the hidden antisymmetric tensor of quantum mechanics. *Class. Quant. Grav.*, 2000, v. 17, 3965–4006. arXiv: hep-th/9909201.

16. Porrier B. Reconciling semiclassical and Bohmian mechanics. I. Stationary states. *J. Chem. Phys.*, 2004, v. 121, 4501–15.

17. Wyatt R. E. Quantum Dynamics with Trajectories. Springer, New York, 2005, pp. 354–68.

18. Floyd E. R. A trajectory interpretation of transmission and reflection. *Phys. Essays.*, 1994, v. 7, 135–145.

19. Floyd E. R. A trajectory interpretation of tunneling. *An. Fond. L. de Broglie*, 1995, v. 20, 263–79.

20. Floyd E. R. The form the normal mode that ensures escape from a surface channel. Proceedings 16th International Congress of Acoustics and 135th Meeting Acoustical Society of America, v. II. Kuhl P. K. and Crum L. A., eds. Acoustical Society of America, Woodbury, NY, 1998, pp. 951–2.

21. Floyd E. R. Interference, reduced action and trajectories. *Found. Phys.*, 2007, v. 37, 1386–1402. arXiv: quant-ph/0605120v3.

22. Pandey A., Porier B., Peralta L., Siddique M., Ho Y.-C., and Farooq H. An unorthodox study of bidirectional light waves. 2019 Joint Meeting of Texas Sections of APS, APT and Zone 13 of SPS, v. 64 Number 18, E01.00005, 2019.

23. Faraggi A. E. and Matone M. The Geometrical Origins of Dark Energy. arXiv: 2006.11935.

24. An F. P., Balantekin A. B., Band H. R., Bishai M., (Daya Bay Collaboration). Measurement of the reactor neutrino flux and spectrum at Daya Bay. *Phys. Rev. Lett.*, 2016, v. 116, 061801. Erratum: *Phys. Rev. Lett.*, 2017, v. 118, 099902.

25. Faraggi A. E. OPERA data and the equivalence principal of quantum mechanics. *Eur. Phys. J.*, 2011, v 72, 1944. arXiv: 1110.1857v2.

26. Arafune A. and Takeda G. Total Reflection of Relic Neutrinos from Material Targets. U. of Tokyo, ICEPP Report, ut-icepp 08-02, unpublished.

27. Dwight H. B. Table of Integrals and Other Mathematical Data. MacMillan, New York, 1961. p. 82, ¶ 403.4 & ¶ 403.5.

28. Pontecorvo B. Inverse beta decay and nonobservation of lepton charge. *Sov. Phys. JETP*, 1958, v. 7, 172–3; in Russian: *Zh. Eksp. Teor. Fiz.*, 1958, v. 4, 247–9.

29. Perelman C. C. Bohm's potential classical/quantum duality and repulsive gravity. *Phys. Lett. B*, 2019, v. 778, 546–51.

30. Bohm D. A suggested interpretation of quantum theory in terms of "Hidden Variables". *Phys. Rev.*, 1953, v. 85, 166–79.

31. Aguilar-Arcvalo A. A., Brown B. C., Bugel L., Cheng G., (MiniBooNE Collaboration). Observation of a significant excess of events in the MiniBooNE short-baseline neutrino experiment. arXiv: 1805.12028.

32. Cho A. Report of sterile neutrino resurrection may be greatly exaggerated. *Science*, 10.1126/science.aau3773, 2018. url: <http//www.sciencemag.org/news/2018/06/reports-sterile-neutrino-s-resurrection-may-be-greatly-exaggerated> (accessed 20 August 2020).

33. Floyd E. R. The Ermakov invariant for the trajectory representation of quantum mechanics. *Phys. Lett. A*, 1996, v. 214, 259–65.