Issue 2 2020 | Volume 16

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

ISSN 1555-5534

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

# PROGRESS IN PHYSICS

**A quarterly issue scientific journal, registered with the Library of Congress (DC, USA). This journal is peer reviewed and included in the abstracting and indexing coverage of: Mathematical Reviews and MathSciNet (AMS, USA), DOAJ of Lund University (Sweden), Scientific Commons of the University of St. Gallen (Switzerland), Open-J-Gate (India), Referativnyi Zhurnal VINITI (Russia), etc.**

Electronic version of this journal:
http://www.ptep-online.com



Postal Address

Department of Mathematics and Science,
University of New Mexico,
705 Gurley Ave., Gallup, NM 87301, USA



**ISSN:** 1555-5534 (print)
**ISSN:** 1555-5615 (online)

**Standard Address Number: 297-5092**
**Printed in the United States of America**

**October 2020** **Vol. 16, Issue 2**

## 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:

$$\frac{d^2x^\lambda}{ds^2} - \Gamma^\lambda_{\alpha\delta}\frac{dx^\alpha}{ds}\frac{dx^\delta}{ds} = 0 \tag{1}$$

where $ds = cd\tau$ is the line element, $\tau$ is the relativistic proper time, $c = 2.99792458 \times 10^8\,\mathrm{m\,s^{-1}}$ (CODATA 2018) is the speed of light in *vacuo*, $x^\mu$ 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 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 ($\overline{\Gamma}^\lambda_{\mu\nu}$) 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

*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 $(x, y, z)$ rectungular system of coordinates to say the spherical $(r, \theta, \varphi)$, the resulting affine $\Gamma^{\lambda'}_{\mu'\nu'}$ is not zero – i.e. $\Gamma^{\lambda'}_{\mu'\nu'} \neq 0$.

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 Schwarzchild 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:

$$g_{\mu\nu;\alpha} = g_{\mu\nu,\alpha} - \Gamma^{\delta}_{\alpha\mu} g_{\delta\nu} - \Gamma^{\delta}_{\alpha\nu} g_{\mu\delta} = 0 \tag{2}$$

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.:

$$g_{\mu\nu;\alpha} = g_{\mu\nu,\alpha} - \Gamma^{\delta}_{\alpha\mu} g_{\delta\nu} - \Gamma^{\delta}_{\alpha\nu} g_{\mu\delta} = 0\,, \tag{3}$$

$$g_{\mu\nu;\alpha} = g_{\mu\nu,\alpha} - \Gamma^{\delta}_{\alpha\mu} g_{\delta\nu} - \Gamma^{\delta}_{\alpha\nu} g_{\mu\delta} = 0\,. \tag{4}$$

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:

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}\, g^{\delta\lambda}\left(g_{\delta\mu,\nu} + g_{\nu\delta,\mu} - g_{\mu\nu,\delta}\right). \tag{5}$$

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:

$$\Gamma^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^{\delta}} \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} \Gamma^{\delta}_{\mu\nu} + \underbrace{\frac{\partial x^{\lambda'}}{\partial x^{\delta}} \frac{\partial^2 x^{\delta}}{\partial x^{\mu'} \partial x^{\nu'}}}_{\textbf{extra term}} . \qquad (6)$$

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:

$$\overbrace{\underbrace{R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu}}_{\textit{beautiful and splendour}}}^{\textbf{MARBLE}} = \overbrace{\underbrace{\kappa_E T_{\mu\nu}}_{\textit{ugly and loathsome}}}^{\textbf{WOOD}} \qquad (7)$$

where $\kappa_E = 8\pi G/c^4$ is Einstein's constant of gravitation, $G = 6.67430(15) \times 10^{-11}\,\mathrm{kg^{-1}\,m^3\,s^{-2}}$ (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 $\varrho_g$ is the density of matter, $p$ is the pressure, $v_\mu$ the four-velocity, and $\Lambda$ 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 $R_{\mu\nu} - R g_{\mu\nu}/2$ and material tensor $T_{\mu\nu}$) *via* some mere constant $\kappa_E$. 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 $\boldsymbol{E}$ and $\boldsymbol{B}$-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 $(\boldsymbol{X}, \boldsymbol{Y})$ 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 $(\boldsymbol{E})$ and magnetic $(\boldsymbol{B})$ fields and the associated field equations with Maxwell's equations [14], i.e.: $\boldsymbol{X} := \boldsymbol{E}, \boldsymbol{Y} := \boldsymbol{B}$.

In brief, what Heras [43] proved is that if $\varrho_{\mathrm{e}}$ is the charge density and $\vec{J}$ is the associated current corresponding to this charge, i.e.:

$$\frac{\partial \varrho_{\mathrm{e}}}{\partial t} = -\vec{\nabla} \cdot \vec{J}, \tag{8}$$

then, there must exist two corresponding fields, $\boldsymbol{X}$ and $\boldsymbol{Y}$, that satisfy the following set of equations:

$$\begin{aligned} \vec{\nabla} \cdot \boldsymbol{X} &= \alpha \varrho_{\mathrm{e}} && \text{(a)} \\ \vec{\nabla} \cdot \boldsymbol{Y} &= 0 && \text{(b)} \\ \vec{\nabla} \times \boldsymbol{X} + \gamma \frac{\partial \boldsymbol{Y}}{\partial t} &= 0 && \text{(c)} \\ \vec{\nabla} \times \boldsymbol{Y} - \frac{\beta}{\alpha} \frac{\partial \boldsymbol{X}}{\partial t} &= \beta \vec{J} && \text{(d)} \end{aligned} \tag{9}$$

where $\alpha$, $\beta$, $\gamma$ 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 $\boldsymbol{X}$ and $\boldsymbol{Y}$ 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 $\varrho_{\mathrm{g}}$ has an associated vector field. We shall denote this vector field by the symbol $\boldsymbol{A}_{\mathrm{g}}$ 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 $\varrho_{\mathrm{g}}$ with $\boldsymbol{A}_{\mathrm{g}}$ has been known for a considerable amount of time now. That is, for more than a century ($\geq$ 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:

$$\vec{\nabla} \cdot \vec{\mathfrak{g}} = -4\pi G \varrho_{\mathrm{g}} \tag{10}$$

where $\vec{\mathfrak{g}}$ is the gravitational field intensity at a given point in the gravitational field, $\varrho_{\mathrm{g}}$ is the gravitational potential, and from the equation of conservation of mass-energy and momentum, namely: $\partial \varrho_{\mathrm{g}}/\partial t = -\vec{\nabla} \cdot \vec{\mathfrak{J}}$, where $\vec{\mathfrak{J}} = \varrho_{\mathrm{g}} \boldsymbol{v}$, is the momentum density with $\boldsymbol{v}$ being the velocity of the material whose density is $\varrho_{\mathrm{g}}$.

In order to see this, from (10) we know very well that: $\dot{\varrho}_{\mathrm{g}} = -(1/4\pi G)(\vec{\nabla} \cdot \dot{\vec{\mathfrak{g}}})$. Let us set: $\widetilde{\mu} = 1/4\pi G$, so that: $\dot{\varrho}_{\mathrm{g}} = (1/4\pi G)(\vec{\nabla} \cdot \dot{\vec{\mathfrak{g}}})$ can now be written as: $\dot{\varrho}_{\mathrm{g}} = -\widetilde{\mu}\vec{\nabla} \cdot \dot{\vec{\mathfrak{g}}}$. From this, it follows that:

$$\frac{\partial \varrho_{\mathrm{g}}}{\partial t} = -\vec{\nabla} \cdot \vec{\mathfrak{J}} = -\vec{\nabla} \cdot \left( \widetilde{\varepsilon} \frac{\partial \vec{\mathfrak{g}}}{\partial t} \right), \tag{11}$$

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 ($\mu_0$) and permittivity ($\epsilon_0$) 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 $\partial\varrho_{\mathrm{g}}/\partial t=-\vec{\nabla}\cdot\vec{\mathfrak{J}}$ 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:

$$\begin{aligned}
\vec{\nabla}\cdot\vec{\mathfrak{g}} &= -\varrho_{\mathrm{g}}/\widetilde{\varepsilon} && \text{(a)}\\
\vec{\nabla}\times\vec{\mathfrak{g}} &= -\frac{1}{\widetilde{c}}\frac{\partial\vec{\mathfrak{B}}}{\partial t} && \text{(b)}\\
\vec{\nabla}\cdot\vec{\mathfrak{B}} &= 0 && \text{(c)}\\
\vec{\nabla}\times\vec{\mathfrak{B}} &= -\widetilde{\mu}\vec{\mathfrak{J}}+\frac{1}{\widetilde{c}^2}\frac{\partial\vec{\mathfrak{g}}}{\partial t}. && \text{(d)}
\end{aligned} \tag{15}$$

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 $\mathrm{g}_{\mu\nu}\mapsto e^{2\phi}\mathrm{g}_{\mu\nu}$, Weyl [1] built his geometry by supplementing the Christoffel affine connection $\Gamma^{\lambda}_{\mu\nu}$ of Riemann geometry with a tensorial affine $\mathcal{W}^{\lambda}_{\mu\nu}$:

$$\mathcal{W}^{\lambda}_{\mu\nu}=\mathrm{g}^{\lambda}{}_{\mu}\mathcal{A}_{\nu}+\mathrm{g}^{\lambda}{}_{\nu}\mathcal{A}_{\mu}-\mathrm{g}_{\mu\nu}\mathcal{A}^{\lambda}\,, \tag{16}$$

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 $\mathrm{g}_{\mu\nu}$ and the four-vector $\mathcal{A}_{\mu}$:

$$\begin{aligned}
\mathrm{g}_{\mu\nu} &\longmapsto e^{2\chi}\mathrm{g}_{\mu\nu}\\
\mathcal{A}_{\mu} &\longmapsto \mathcal{A}_{\mu}+\kappa_0^{-1}\partial_{\mu}\chi
\end{aligned}\,, \tag{18}$$

where $\chi=\chi(\boldsymbol{r},t)$ is a well behaved, arbitrary, smooth, differentiable, integrable and uniform continuous scalar function, and $\kappa_0$ is a constant with the dimensions of inverse length. This constant $\kappa_0$ has been introduced for the purposes of dimensional consistency, since we here assume that the four-vector $\mathcal{A}_{\mu}$ and the true scalar $\chi$ 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

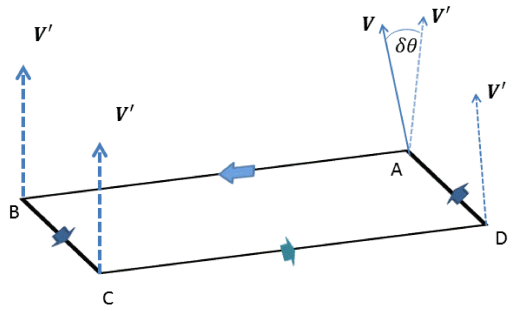



Fig. 1: **Parallel Transport:** The vector $V$ is parallel transported in a closed circuit. Upon arrival at its original position, the vector is not equal to the original vector and this is a result of the curvature of the space in question.

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 $\boldsymbol{V}$ is transported in a closed circuit such that it returns to its original position and $\boldsymbol{V}'$ is the resulting vector after parallel transport; in normal Riemann geometry, while $|\boldsymbol{V}| = |\boldsymbol{V}'|$, the angle $\delta\theta$ between these two vectors, while it can in some cases equal zero, is not necessarily zero i.e. $\boldsymbol{V} \cdot \boldsymbol{V}' \neq 0$. However, on the RWS, we have for all points of space and time on this spacetime the constraints $|\boldsymbol{V}| = |\boldsymbol{V}'|$ and $\boldsymbol{V} \cdot \boldsymbol{V}' = 0$: 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 $\mathrm{g}_{\mu\nu}$ to $e^{2\chi}\mathrm{g}_{\mu\nu}$: i.e. after the transformation $\mathrm{g}_{\mu\nu} \longmapsto e^{2\chi}\mathrm{g}_{\mu\nu}$, 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 $\mathcal{W}^\lambda_{\mu\nu}$ to the Christoffel three-symbol $\Gamma^\lambda_{\mu\nu}$, that is to say, if $\widetilde{\Gamma}^\lambda_{\mu\nu}$ is the new Christoffel symbol for the Weyl space, then:

$$\widetilde{\Gamma}^\lambda_{\mu\nu} = \Gamma^\lambda_{\mu\nu} + \mathcal{W}^\lambda_{\mu\nu}\,. \tag{19}$$

Because $\mathcal{W}^\lambda_{\mu\nu}$ is a tensor, the fundamental transformational properties of the new Christoffel three-symbol $\widetilde{\Gamma}^\lambda_{\mu\nu}$ are the same as the old Christoffel three-symbol $\Gamma^\lambda_{\mu\nu}$; 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 $\overline{\mathrm{g}}_{\mu\nu}$, instead of making it conformal only at the instance of a gauge transformation, we chose that it $(\overline{\mathrm{g}}_{\mu\nu})$ be intrinsically and inherently conformal. That is to say, the fundamental metric $\overline{\mathrm{g}}_{\mu\nu}$ of the RWS be such that $\overline{\mathrm{g}}_{\mu\nu} = \varrho\,\mathrm{g}_{\mu\nu}$, where $\mathrm{g}_{\mu\nu}$ 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 $\varrho$ is a scalar. However in [33], this function takes a decisive new role: ... it (the scalar $\chi$) 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 $x^\lambda$ i.e.:

$$\mathcal{V}_\mu = \frac{\partial_\mu \mathcal{S}}{\mathcal{S}} = \partial_\mu \ln \mathcal{S} \tag{20}$$

where:

$$\mathcal{S} = \mathcal{B}_\delta x^\delta \,. \tag{21}$$

From (20) and (21), it follows that:

$$\mathcal{V}_\mu = \mathcal{B}_\mu + \frac{x^\delta \partial_\mu \mathcal{B}_\delta}{\mathcal{S}} \,. \tag{22}$$

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:

$$\mathcal{V}_{\mu'} = \frac{\partial x^\delta}{\partial x^{\mu'}} \mathcal{V}_\delta + \frac{\partial^2 x^\delta}{\partial x^{\mu'} \partial x^{\Omega'}} \,. \tag{23}$$

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 $\mathrm{g}_{\mu\nu;\sigma} = 0$ of Riemann geometry. As already alluded, the condition $\mathrm{g}_{\mu\nu;\sigma} = 0$ is the foundation stone of Riemann geometry. We will uphold this covariant derivative condition under the Weyl conformal transformation $\mathrm{g}_{\mu\nu} \longmapsto \overline{\mathrm{g}}_{\mu\nu} = \varrho \mathrm{g}_{\mu\nu}$ of the metric i.e. $\overline{\mathrm{g}}_{\mu\nu;\sigma} = 0$. Likewise, the condition $\overline{\mathrm{g}}_{\mu\nu;\sigma} = 0$ is to be taken as the foundation stone of the new hybrid Riemann-Weyl geometry. Written in full, the equation $\overline{\mathrm{g}}_{\mu\nu;\sigma} = 0$ is given by:

$$\overline{\mathrm{g}}_{\mu\nu;\sigma} = \varrho \left[ \mathrm{g}_{\mu\nu,\sigma} + \mathrm{g}_{\mu\nu} \left( \frac{\partial_\sigma \varrho}{\varrho} \right) - \mathrm{g}_{\mu\delta} \overline{\Gamma}^\delta_{\nu\sigma} - \mathrm{g}_{\delta\nu} \overline{\Gamma}^\delta_{\mu\sigma} \right] = 0 \tag{24}$$

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 $\varrho$-quantity as:

$$\varrho = -2 \jmath_\delta x^\delta \tag{25}$$

where $\jmath_\sigma$ is the (gravitational) four-current density. With this definition for $\varrho$, it follows that (24) will reduce to:

$$\mathrm{g}_{\mu\nu,\sigma} - \mathrm{g}_{\mu\delta} \overline{\Gamma}^\delta_{\nu\sigma} - \mathrm{g}_{\delta\nu} \overline{\Gamma}^\delta_{\mu\sigma} = 2 \left( \jmath_\sigma + Q_\sigma \right) \mathrm{g}_{\mu\nu} \tag{26}$$

where $Q_\sigma = x^\delta \partial_\sigma \jmath_\delta / \varrho = x^\delta \partial_\sigma \jmath_\delta / \jmath_\delta x^\delta$. As is the case with Weyl's original geometry [1], the covariant derivative $\mathrm{g}_{\mu\nu;\sigma}$ does not vanish since $\mathrm{g}_{\mu\nu;\sigma} \neq 0$.

#### 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 $\mu, \nu$ and $\sigma$ in $g_{\mu\nu,\sigma}$, i.e.:

$$\begin{aligned} g_{\mu\nu,\sigma} - g_{\mu\delta}\overline{\Gamma}^{\delta}_{\nu\sigma} - g_{\delta\nu}\overline{\Gamma}^{\delta}_{\mu\sigma} &= 2\left(\jmath_{\sigma} + Q_{\sigma}\right) g_{\mu\nu} \quad \text{(a)} \\ g_{\sigma\mu,\nu} - g_{\sigma\delta}\overline{\Gamma}^{\delta}_{\mu\nu} - g_{\delta\nu}\overline{\Gamma}^{\delta}_{\sigma\mu} &= 2\left(\jmath_{\nu} + Q_{\nu}\right) g_{\sigma\mu} \quad \text{(b)} \\ g_{\nu\sigma,\mu} - g_{\nu\delta}\overline{\Gamma}^{\delta}_{\sigma\mu} - g_{\delta\mu}\overline{\Gamma}^{\delta}_{\nu\sigma} &= 2\left(\jmath_{\mu} + Q_{\mu}\right) g_{\nu\sigma} \quad \text{(c)} \end{aligned} \tag{27}$$

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:

$$\begin{aligned} &g_{\mu\nu,\sigma} - g_{\sigma\mu,\nu} - g_{\nu\sigma,\mu} + g_{\sigma\delta}\left[\overline{\Gamma}^{\delta}_{\mu\nu} + g_{\delta\nu}\overline{\Gamma}^{\delta}_{\sigma\mu}\right] = \\ &= 2\left[\left(\jmath_{\sigma} + Q_{\sigma}\right) g_{\mu\nu} - \left(\jmath_{\nu} + Q_{\nu}\right) g_{\sigma\mu} - \left(\jmath_{\mu} + Q_{\mu}\right) g_{\nu\sigma}\right]. \end{aligned} \tag{28}$$

Contracting the $\sigma$-index in (28) by multiplying (28) throughout by $g^{\sigma\lambda}$ and thereafter resetting this $\sigma$-index to $\delta$, we obtain:

$$\begin{aligned} &- g^{\delta\lambda}\left[g_{\delta\mu,\nu} + g_{\nu\delta,\mu} - g_{\mu\nu,\delta}\right] + 2\overline{\Gamma}^{\lambda}_{\mu\nu} = \\ &- 2g^{\delta\lambda}\left[\left(\jmath_{\nu} + Q_{\nu}\right) g_{\delta\mu} + \left(\jmath_{\mu} + Q_{\mu}\right) g_{\nu\delta} - \left(\jmath_{\delta} + Q_{\delta}\right) g_{\mu\nu}\right], \end{aligned} \tag{29}$$

hence:

$$\overline{\Gamma}^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\mu\nu} - \mathcal{W}^{\lambda}_{\mu\nu} - Q^{\lambda}_{\mu\nu}, \tag{30}$$

where $\Gamma^{\lambda}_{\mu\nu}$ is the usual Christoffel three-symbol given in (5), and

$$\mathcal{W}^{\lambda}_{\mu\nu} = g^{\lambda}{}_{\mu}\jmath_{\nu} + g^{\lambda}{}_{\nu}\jmath_{\mu} - g_{\mu\nu}\jmath^{\lambda} \tag{31}$$

is the (redefined) Weyl tensor, and the new non-tensorial object:

$$Q^{\lambda}_{\mu\nu} = g^{\lambda}{}_{\mu}Q_{\nu} + g^{\lambda}{}_{\nu}Q_{\mu} - g_{\mu\nu}Q^{\lambda} \tag{32}$$

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:

$$\Gamma^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^{\delta}}\frac{\partial x^{\mu}}{\partial x^{\mu'}}\frac{\partial x^{\nu}}{\partial x^{\nu'}}\Gamma^{\delta}_{\mu\nu} + \frac{\partial x^{\lambda'}}{\partial x^{\delta}}\frac{\partial^2 x^{\delta}}{\partial x^{\mu'}\partial x^{\nu'}} \tag{33}$$

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:

$$Q^{\lambda'}_{\mu'\nu'} = \frac{\partial x^{\lambda'}}{\partial x^{\delta}}\frac{\partial x^{\mu}}{\partial x^{\mu'}}\frac{\partial x^{\nu}}{\partial x^{\nu'}}Q^{\delta}_{\mu\nu} + \frac{\partial x^{\lambda'}}{\partial x^{\delta}}\frac{\partial^2 x^{\delta}}{\partial x^{\mu'}\partial x^{\nu'}}, \tag{34}$$

then it follows and goes without saying that the object $\overline{\Gamma}^{\lambda}_{\mu\nu}$ will clearly be a tensor. Because $Q_{\mu}$ 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 $Q_{\mu}$ 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 $v^\lambda$ 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:

$$\overline{R}^\lambda_{\mu\sigma\nu} = \overbrace{\overline{\Gamma}^\lambda_{\mu\nu,\sigma} - \overline{\Gamma}^\lambda_{\mu\sigma,\nu}}^{\textbf{linear terms}} + \underbrace{\overline{\Gamma}^\lambda_{\delta\sigma}\overline{\Gamma}^\delta_{\mu\nu} - \overline{\Gamma}^\lambda_{\delta\nu}\overline{\Gamma}^\delta_{\mu\sigma}}_{\textbf{non-linear terms}} \,. \tag{36}$$

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 (†) 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:

$$\begin{aligned} g_{\mu\nu} &= \mathcal{A}_\mu\mathcal{A}_\nu \,, \qquad & g_\mu{}^\nu &= \mathcal{A}_\mu\mathcal{A}^\nu \,, \\ g^{\mu\nu} &= \mathcal{A}^\mu\mathcal{A}^\nu \,, \qquad & g^\mu{}_\nu &= \mathcal{A}^\mu\mathcal{A}_\nu \,. \end{aligned} \tag{40}$$

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 $4\times4$ 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.:

$$\begin{aligned} V_\mu &= g_{\mu\delta}V^\delta = g_\mu{}^\delta V_\delta \\ V^\mu &= g^{\mu\delta}V_\delta = g^\mu{}_\delta V^\delta \end{aligned} \quad . \tag{43}$$

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 $\overline{R}^\lambda_{\mu\sigma\nu}$ vanishes at all points of this spacetime, i.e.:

$$\overline{R}^\lambda_{\mu\sigma\nu} = 0\,. \qquad \textbf{(natural gauge)} \tag{44}$$

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 $\overline{\Gamma}^\lambda_{\mu\nu}$ is non-vanishing, i.e. $\overline{\Gamma}^\lambda_{\mu\nu} \neq 0$. 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 $\overline{R}^\lambda_{\mu\sigma\nu}$ 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 $\overline{R}^\lambda_{\mu\sigma\nu}$, 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 $\overline{R}^\lambda_{\mu\sigma\nu}$ as is given in (38) as follows:

$$\overline{R}^\lambda_{\mu\sigma\nu} = R^\lambda_{\mu\sigma\nu} - \underbrace{\left(\mathcal{W}^\lambda_{\mu\sigma\nu} + Q^\lambda_{\mu\sigma\nu}\right)}_{\mathcal{T}^\lambda_{\mu\sigma\nu}} \tag{45}$$

where:

$$\begin{aligned} R^\lambda_{\mu\sigma\nu} &= \Gamma^\lambda_{\mu\nu,\sigma} - \Gamma^\lambda_{\mu\sigma,\nu} && \text{(a)} \\ \mathcal{W}^\lambda_{\mu\sigma\nu} &= \mathcal{W}^\lambda_{\mu\nu,\sigma} - \mathcal{W}^\lambda_{\mu\sigma,\nu} && \text{(b)} \\ Q^\lambda_{\mu\sigma\nu} &= Q^\lambda_{\mu\nu,\sigma} - Q^\lambda_{\mu\sigma,\nu} && \text{(c)} \end{aligned} \tag{46}$$

are the *linear Riemann curvature tensor* (46*a*), the *linear Weyl curvature tensor* (46*b*), and the *linear $Q$-curvature tensor* (46*c*) 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 $\varrho$, i.e. $\mathcal{W}^\lambda_{\mu\sigma\nu} = \mathcal{W}^\lambda_{\mu\sigma\nu}(\mathcal{J}_\alpha)$ and $Q^\lambda_{\mu\sigma\nu} = Q^\lambda_{\mu\sigma\nu}(\varrho)$. The $Q$-tensor is a direct function of $\varrho$ while the Weyl tensor is not – remember (25) that $\mathcal{J}_\alpha = -\frac{1}{2}\partial_\alpha\varrho$, 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 $\varrho$ 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:

$$\mathcal{T}^\lambda_{\mu\sigma\nu} = \mathcal{W}^\lambda_{\mu\sigma\nu} + Q^\lambda_{\mu\sigma\nu}\,. \tag{47}$$

With the above definition (47) of the material tensor, it follows that the Riemann-Weyl curvature tensor $\overline{R}^\lambda_{\mu\sigma\nu}$ 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:

$$\overline{R}^\lambda_{\mu\sigma\nu} = R^\lambda_{\mu\sigma\nu} - \mathcal{T}^\lambda_{\mu\sigma\nu}\,. \tag{48}$$

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:

$$R^\lambda_{\mu\sigma\nu} = \mathcal{T}^\lambda_{\mu\sigma\nu}\,. \tag{49}$$

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 ($\mathcal{W}^{\lambda}_{\upsilon\mu\sigma,\nu}$; $\mathcal{Q}^{\lambda}_{\upsilon\mu\sigma,\nu}$) of this tensor also satisfy the Voss-Bianchi identities, i.e.:

$$\begin{aligned}
&\mathcal{W}^{\lambda}_{\mu\sigma\nu} + \mathcal{W}^{\lambda}_{\nu\mu\sigma} + \mathcal{W}^{\lambda}_{\sigma\nu\mu} \equiv 0 \quad \text{(a)}\\
&\mathcal{Q}^{\lambda}_{\mu\sigma\nu} + \mathcal{Q}^{\lambda}_{\nu\mu\sigma} + \mathcal{Q}^{\lambda}_{\sigma\nu\mu} \equiv 0 \quad \text{(b)}\\
&\mathcal{W}^{\lambda}_{\upsilon\mu\sigma,\nu} + \mathcal{W}^{\lambda}_{\upsilon\nu\mu,\sigma} + \mathcal{W}^{\lambda}_{\upsilon\sigma\nu,\mu} \equiv 0 \quad \text{(c)}\\
&\mathcal{Q}^{\lambda}_{\upsilon\mu\sigma,\nu} + \mathcal{Q}^{\lambda}_{\upsilon\nu\mu,\sigma} + \mathcal{Q}^{\lambda}_{\upsilon\sigma\nu,\mu} \equiv 0 \quad \text{(d)}
\end{aligned} \tag{54}$$

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 $\mathcal{Q}^{\lambda}_{\upsilon\mu\sigma,\nu}$ 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 $2\Gamma^{\lambda}_{\mu\nu} = \mathrm{g}^{\delta\lambda}\left(\mathrm{g}_{\delta\mu,\nu} + \mathrm{g}_{\nu\delta,\mu} - \mathrm{g}_{\mu\nu,\delta}\right)$, and from the present new findings that the decomposed Riemann metric tensor is such that $\mathrm{g}_{\mu\nu} = \mathcal{A}_{\mu}\mathcal{A}_{\nu}$. 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:

$$\begin{aligned}
2\Gamma^{\lambda}_{\mu\nu} = \mathrm{g}^{\delta\lambda}\Big[ &\overbrace{\mathcal{A}_{\delta}\mathcal{A}_{\mu,\nu}}^{\textbf{Term I}} + \overbrace{\mathcal{A}_{\mu}\mathcal{A}_{\delta,\nu}}^{\textbf{Term II}} + \overbrace{\mathcal{A}_{\delta}\mathcal{A}_{\nu,\mu}}^{\textbf{Term III}} +\\
&+ \underbrace{\mathcal{A}_{\nu}\mathcal{A}_{\delta,\mu}}_{\textbf{Term IV}} - \underbrace{\mathcal{A}_{\mu}\mathcal{A}_{\nu,\delta}}_{\textbf{Term V}} - \underbrace{\mathcal{A}_{\nu}\mathcal{A}_{\mu,\delta}}_{\textbf{Term VI}} \Big].
\end{aligned} \tag{55}$$

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:
$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} \mathrm{g}^{\delta\lambda} \left[\mathcal{A}_\mu \mathfrak{F}_{\delta\nu} + \mathcal{A}_\nu \mathfrak{F}_{\delta\mu} + \mathcal{A}_\delta \Omega_{\mu\nu}\right]. \tag{62}$$

Now, multiplying the terms in the square bracket by $\mathrm{g}^{\delta\lambda}$, the meaning of which is that we have to raise the $\delta$-index in these square brackets and reset it so that it now equals $\lambda$, 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:
$$\Gamma^\lambda_{\mu\nu} = \breve{\Gamma}^\lambda_{\mu\nu} + \Omega^\lambda_{\mu\nu} \tag{64}$$
where:
$$\breve{\Gamma}^\lambda_{\mu\nu} = \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda{}_\nu + \mathcal{A}_\nu \mathfrak{F}^\lambda{}_\mu\right) \tag{65}$$
and:
$$\Omega^\lambda_{\mu\nu} = \frac{1}{2} \mathcal{A}^\lambda \Omega_{\mu\nu}. \tag{66}$$

The object $\breve{\Gamma}^\lambda_{\mu\nu}$ 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 (46*a*); so doing, we obtain:
$$\begin{aligned} R^\lambda_{\mu\sigma\nu} &= \breve{\Gamma}^\lambda_{\mu\nu,\sigma} - \breve{\Gamma}^\lambda_{\mu\sigma,\nu} + \Omega^\lambda_{\mu\nu,\sigma} - \Omega^\lambda_{\mu\sigma,\nu} \\ &= \breve{R}^\lambda_{\mu\sigma\nu} + \Omega^\lambda_{\mu\sigma\nu} \end{aligned} \tag{68}$$
where:
$$\begin{aligned} \breve{R}^\lambda_{\mu\sigma\nu} &= \breve{\Gamma}^\lambda_{\mu\nu,\sigma} - \breve{\Gamma}^\lambda_{\mu\sigma,\nu} \\ \Omega^\lambda_{\mu\sigma\nu} &= \Omega^\lambda_{\mu\nu,\sigma} - \Omega^\lambda_{\mu\sigma,\nu} \end{aligned} \tag{69}$$
are tensors. The reader will need to verify for themselves that – indeed – these objects are tensors.

Further, we will express $\breve{R}^\lambda_{\mu\sigma\nu}$ in terms of the field tensor $\mathfrak{F}_{\mu\nu}$ by substituting $\breve{\Gamma}^\lambda_{\mu\nu}$ as it is given in (65); so doing, one obtains:
$$\begin{aligned} \breve{R}^\lambda_{\mu\sigma\nu} &= \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda{}_{\nu,\sigma} + \mathcal{A}_\nu \mathfrak{F}^\lambda{}_{\mu,\sigma}\right) - \\ &- \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda{}_{\sigma,\nu} + \mathcal{A}_\sigma \mathfrak{F}^\lambda{}_{\mu,\nu}\right) + \\ &+ \frac{1}{2} \left(\mathcal{A}_{\mu,\sigma} \mathfrak{F}^\lambda{}_\nu + \mathcal{A}_{\nu,\sigma} \mathfrak{F}^\lambda{}_\mu\right) - \\ &- \frac{1}{2} \left(\mathcal{A}_{\mu,\nu} \mathfrak{F}^\lambda{}_\sigma + \mathcal{A}_{\sigma,\nu} \mathfrak{F}^\lambda{}_\mu\right) \\ &= \acute{R}^\lambda_{\mu\sigma\nu} + \grave{R}^\lambda_{\mu\sigma\nu} \end{aligned} \tag{70}$$
where:
$$\begin{aligned} \acute{R}^\lambda_{\mu\sigma\nu} &= \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda{}_{\nu,\sigma} + \mathcal{A}_\nu \mathfrak{F}^\lambda{}_{\mu,\sigma}\right) \\ &- \frac{1}{2} \left(\mathcal{A}_\mu \mathfrak{F}^\lambda{}_{\sigma,\nu} + \mathcal{A}_\sigma \mathfrak{F}^\lambda{}_{\mu,\nu}\right) \end{aligned} \tag{71}$$
and:
$$\begin{aligned} \grave{R}^\lambda_{\mu\sigma\nu} &= \frac{1}{2} \left(\mathcal{A}_{\mu,\sigma} \mathfrak{F}^\lambda{}_\nu + \mathcal{A}_{\nu,\sigma} \mathfrak{F}^\lambda{}_\mu\right) - \\ &- \frac{1}{2} \left(\mathcal{A}_{\mu,\nu} \mathfrak{F}^\lambda{}_\sigma + \mathcal{A}_{\sigma,\nu} \mathfrak{F}^\lambda{}_\mu\right) \end{aligned} \tag{72}$$

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 $\mathrm{g}^{\lambda}{}_{\mu}\mathcal{J}_{\nu}$, we obtain that:

$$\begin{aligned}
\mathcal{T}^{\lambda}_{\mu\sigma\nu} &= \left(\mathrm{g}^{\lambda}{}_{\mu}\mathcal{J}_{\nu,\sigma} + \mathrm{g}_{\nu}{}^{\lambda}\mathcal{J}_{\mu,\sigma} - \mathrm{g}_{\mu\nu}\mathcal{J}^{\lambda}{}_{,\sigma}\right) - \\
&- \left(\mathrm{g}^{\lambda}{}_{\mu}\mathcal{J}_{\sigma,\nu} + \mathrm{g}_{\sigma}{}^{\lambda}\mathcal{J}_{\mu,\nu} - \mathrm{g}_{\mu\sigma}\mathcal{J}^{\lambda}{}_{,\nu}\right) + \\
&+ \left(\mathrm{g}^{\lambda}{}_{\mu,\sigma}\mathcal{J}_{\nu} + \mathrm{g}_{\nu,\sigma}{}^{\lambda}\mathcal{J}_{\mu} - \mathrm{g}_{\mu\nu,\sigma}\mathcal{J}^{\lambda}\right) - \\
&- \left(\mathrm{g}^{\lambda}{}_{\mu,\nu}\mathcal{J}_{\sigma} + \mathrm{g}_{\sigma,\nu}{}^{\lambda}\mathcal{J}_{\mu} - \mathrm{g}_{\mu\sigma,\nu}\mathcal{J}^{\lambda}\right) + \\
&+ Q^{\lambda}_{\mu\sigma\nu} \\
&= \acute{\mathcal{T}}^{\lambda}_{\mu\sigma\nu} + \grave{\mathcal{T}}^{\lambda}_{\mu\sigma\nu} + Q^{\lambda}_{\mu\sigma\nu}
\end{aligned} \tag{74}$$

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:

$$\begin{aligned}
\acute{\mathcal{T}}^{\lambda}_{\mu\sigma\nu} &= \left(\mathrm{g}^{\lambda}{}_{\mu}\mathcal{J}_{\nu,\sigma} + \mathrm{g}_{\nu}{}^{\lambda}\mathcal{J}_{\mu,\sigma} - \mathrm{g}_{\mu\nu}\mathcal{J}^{\lambda}{}_{,\sigma}\right) - \\
&- \left(\mathrm{g}^{\lambda}{}_{\mu}\mathcal{J}_{\sigma,\nu} + \mathrm{g}_{\sigma}{}^{\lambda}\mathcal{J}_{\mu,\nu} - \mathrm{g}_{\mu\sigma}\mathcal{J}^{\lambda}{}_{,\nu}\right) \\
&= \left[\mathrm{g}_{\nu}{}^{\lambda}\mathcal{J}_{\mu,\sigma} - \mathrm{g}_{\mu\nu}\mathcal{J}^{\lambda}{}_{,\sigma}\right] - \left[\mathrm{g}_{\sigma}{}^{\lambda}\mathcal{J}_{\mu,\nu} - \mathrm{g}_{\mu\sigma}\mathcal{J}^{\lambda}{}_{,\nu}\right]
\end{aligned} \tag{75}$$

and:

$$\begin{aligned}
\grave{\mathcal{T}}^{\lambda}_{\mu\sigma\nu} &= \left(\mathrm{g}^{\lambda}{}_{\mu,\sigma}\mathcal{J}_{\nu} + \mathrm{g}_{\nu,\sigma}{}^{\lambda}\mathcal{J}_{\mu} - \mathrm{g}_{\mu\nu,\sigma}\mathcal{J}^{\lambda}\right) - \\
&- \left(\mathrm{g}^{\lambda}{}_{\mu,\nu}\mathcal{J}_{\sigma} + \mathrm{g}_{\sigma,\nu}{}^{\lambda}\mathcal{J}_{\mu} - \mathrm{g}_{\mu\sigma,\nu}\mathcal{J}^{\lambda}\right)\,.
\end{aligned} \tag{76}$$

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.:

$$\overbrace{\underbrace{\acute{R}^{\lambda}_{\mu\sigma\nu}}_{\text{L I}} + \underbrace{\grave{R}^{\lambda}_{\mu\sigma\nu}}_{\text{L II}} + \underbrace{\Omega^{\lambda}_{\mu\sigma\nu}}_{\text{L III}}}^{\text{marble}}_{R^{\lambda}_{\mu\sigma\nu}(\mathcal{A}_{\alpha})} = \overbrace{\underbrace{\acute{\mathcal{T}}^{\lambda}_{\mu\sigma\nu}}_{\text{R I}} + \underbrace{\grave{\mathcal{T}}^{\lambda}_{\mu\sigma\nu}}_{\text{R II}} + \underbrace{Q^{\lambda}_{\mu\sigma\nu}}_{\text{R III}}}^{\text{varnished wood}}_{\mathcal{T}^{\lambda}_{\mu\sigma\nu}(\varrho)}\,. \tag{78}$$

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 $(\alpha,\mu)$ and $(\lambda,\sigma)$-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:

$$\begin{aligned} \mathcal{A}^\beta \acute{\mathcal{T}}^\delta_{\beta\delta\nu} &= -2\mathcal{A}^\delta \jmath_{\delta,\nu} - \jmath^\delta{}_{,\delta}\mathcal{A}_\nu \\ &= -2\mathcal{A}^\delta \partial_\delta \partial_\nu \varrho + (\Box\varrho/2)\,\mathcal{A}_\nu \\ &= -\widetilde{\mu}\mathfrak{J}_\nu + \kappa^2 \mathcal{A}_\nu \end{aligned} \tag{82}$$

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. $\overline{\Gamma}^\lambda_{\mu\nu} = \Gamma^\lambda_{\mu\nu} - \mathcal{W}^\lambda_{\mu\nu} - Q^\lambda_{\mu\nu}$) 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 $\overline{R}^\lambda_{\mu\sigma\nu}$ will vanish identically everywhere, i.e. $\overline{R}^\lambda_{\mu\sigma\nu} \equiv 0$. The equation $\overline{R}^\lambda_{\mu\sigma\nu} \equiv 0$ 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 $\overline{\Gamma}^\delta_{\mu\nu}\overline{\Gamma}^\lambda_{\delta\alpha}$ and $\overline{\Gamma}^\lambda_{\delta\nu}\overline{\Gamma}^\delta_{\mu\alpha}$ in the curvature tensor $\overline{R}^\lambda_{\mu\sigma\nu}$ 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 $\overline{R}^\lambda_{\mu\sigma\nu} \equiv 0$ 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 $F_{\mathrm{g}} = -G\mathcal{M}_{\mathrm{g}} m_{\mathrm{g}}/r^2$ and Coulomb's electrostatic law $F_{\mathrm{e}} = Q\,q/4\pi\epsilon r^2$, 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.:

$$\begin{aligned} \imath\hbar\,\gamma_\mu\partial^\mu\psi &= m_0 c\psi \qquad \text{(a)} \\ \partial^\mu F_{\mu\nu} &= \mu_0 J_\nu \qquad \text{(b)} \\ \partial_\lambda F_{\mu\nu} + \partial_\nu F_{\lambda\mu} + \partial_\mu F_{\nu\lambda} &= 0 \qquad \text{(c)} \end{aligned} \tag{86}$$

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), $\gamma_\mu$, $m_0$, and $\psi$ are the usual four $4\times 4$ 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 $A_\mu$ 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:

$$\begin{aligned} \imath\hbar\,\gamma_\mu^{(a)}\mathcal{A}_\mu\partial^\mu\psi &= m_0 c\psi \qquad \text{(a)} \\ \partial^\mu\mathfrak{F}_{\mu\nu} &= -\widetilde{\mu}\mathfrak{J}_\nu \qquad \text{(b)} \\ \partial_\nu\mathfrak{F}_{\mu\sigma} + \partial_\sigma\mathfrak{F}_{\nu\mu} + \partial_\mu\mathfrak{F}_{\sigma\nu} &\equiv 0\,. \qquad \text{(c)} \end{aligned} \tag{87}$$

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).

### Acknowledgements

We are grateful to the *Editorial Board of the Progress in Physics Journal* for the 100% waiver on the page charges on all our earlier publications in the present journal. Publication of the present reading has been made possible by *Prof. Dr. Sohan Jheeta* – we are grateful for his kind assistance.

Received on April 5, 2020

### 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 ($\Delta p$) and position ($\Delta x$) 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 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

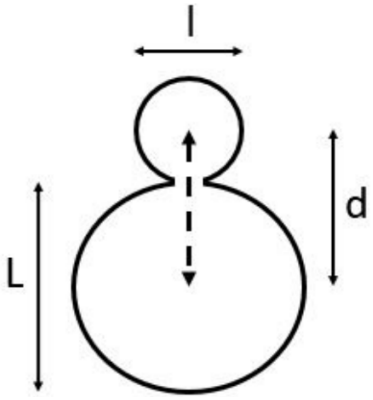



Fig. 1: The asymmetrical metal cavity. A photon moves back and forth along the dashed arrow.

$$\Delta p_w = \frac{\hbar}{2L} \tag{2}$$

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 $\rho$ is the density of the metal walls and $\delta$ 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 $a$

$$L = \sqrt[4]{\frac{8\hbar c}{15\pi\rho\delta a}}\,. \tag{7}$$

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

$$L = 129\,\text{nm}\,. \tag{8}$$

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 nanostructures 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**

Thank you to M. Fiddy and J. Lucio for useful discussions and to DARPA grant HR001118C0125.

Submitted on May 20, 2020

## 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

$$\langle Q_i \rangle_t = \mathrm{Tr}(Q_i\, \rho(t)) \tag{2}$$

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 $\lambda_i$ [11]. Physically speaking, the possible results of measurement correspond to the eigenvalues of $Q$, whereas the probability of getting the result $\lambda_i$ (upon measuring the state $\rho$) 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

$$\begin{aligned} q_{ij} + i r_{ij} = \frac{\mathrm{Re}\; a_{ij} + \mathrm{Re}\; a_{ji}}{2} + i\,\frac{\mathrm{Im}\; a_{ij} - \mathrm{Im}\; a_{ji}}{2} + \\ + i\,\frac{\mathrm{Im}\; a_{ij} + \mathrm{Im}\; a_{ji}}{2} + \frac{\mathrm{Re}\; a_{ij} - \mathrm{Re}\; a_{ji}}{2} = a_{ij}, \end{aligned} \tag{10}$$

which implies that

$$A = Q + i\,R. \tag{11}$$

□

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.

$$A \to (Q_1, Q_2), \text{ where } Q_1, Q_2 \in B_*(\mathcal{H}). \tag{12}$$

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, $\mu$ 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:*

$$K_\mu(\mathbb{L}, Q_i) := \mathrm{Span}\{Q_i, \mathbb{L}^*Q_i, (\mathbb{L}^*)^2Q_i, ..., (\mathbb{L}^*)^{\mu-1}Q_i\}. \tag{18}$$

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 $\mathbb{M}_N(\mathbb{C})$. □

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.

Received on July 6, 2020

## 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.

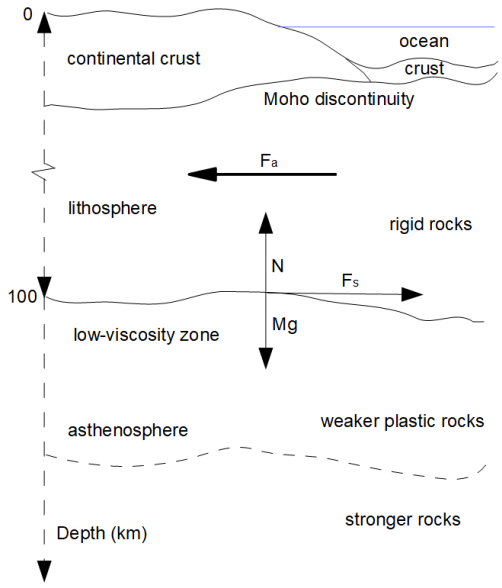
Given the structure of the lithosphere and the plastic and viscous nature of the upper part of the asthenosphere (low-viscosity zone LVZ [5, pp. 11,181]), it is quite conceivable that the lithosphere could move over the asthenosphere by a slip displacement movement, given the appropriate applied force 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 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.

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.




### 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 $\mu_s$. 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 $58.8\times10^6\,\mathrm{km}^2$ [4, p. 7]. For an average thickness ~100 km and an average density $\rho \sim 3.3\,\mathrm{gm/cm}^3$ (see section 1), the mass of the North American plate is given by $M = 1.8\times10^{22}\,\mathrm{kg}$. Using these values and an estimated static coefficient of friction of 0.28 (greasy nickel) [8], the slip condition (2) then becomes

$$F_a > 5\times10^{22}\,\mathrm{N}\,, \tag{3}$$

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. $F_a > 10^{23}\,\mathrm{N}$). 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 $\mu_k$ is smaller than the static coefficient of friction $\mu_s$. 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 $a = -1.2\,\mathrm{m\,s}^{-2}$. 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. $a = -2.4\,\mathrm{m\,s}^{-2}$). 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 $\theta$ with respect to the surface of the plate. For a perpendicular angle of incidence $\theta = 90^\mathrm{o}$, 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 $\Delta t_c$, 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 $d \sim 20\,\mathrm{km}$, mass $m \sim 2\times10^{16}\,\mathrm{kg}$, $v \sim 30\,\mathrm{km/s}$, and use an angle of incidence $\theta = 45^\mathrm{o}$ [12, 13]. Then substituting into (8) and using (3), we obtain slip condition

$$\frac{4\times10^{20}}{\Delta t_p} > 5\times10^{22}\,\mathrm{N}\,, \tag{9}$$

which is dependent on the momentum transfer time. We consider three momentum transfer times: 1 s, 1 ms and 1 $\mu$s:

$$\begin{aligned} &\text{for } \Delta t_p = 1\,\text{s}, && 4\times10^{20}\,\text{N} \not> 5\times10^{22}\,\text{N}\,,\\ &\text{for } \Delta t_p = 1\,\text{ms}, && 4\times10^{23}\,\text{N} > 5\times10^{22}\,\text{N}\,,\\ &\text{for } \Delta t_p = 1\,\mu\text{s}, && 4\times10^{26}\,\text{N} > 5\times10^{22}\,\text{N}\,. \end{aligned} \tag{10}$$

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 $\mu$s), 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

$$E_i = \frac{1}{2}\,mv^2\,, \tag{11}$$

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

$$E_f = \frac{1}{2}\,(M+m)\,V^2 + E_{rel}\,, \tag{12}$$

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

$$\frac{1}{2}\,(M+m)\,V^2 = \frac{1}{2}\,\epsilon\,mv^2\,, \tag{13}$$

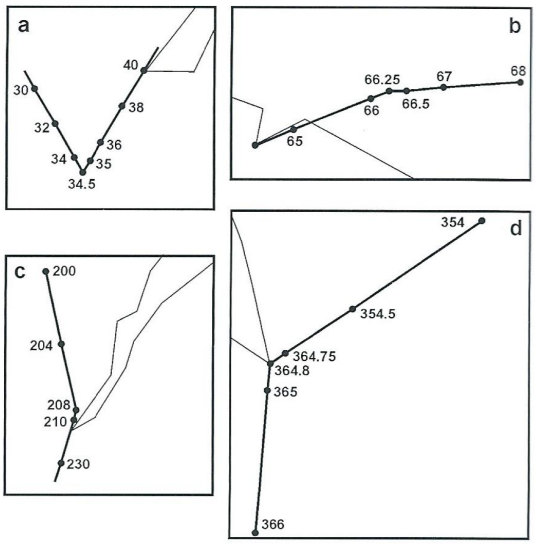



Fig. 2: Figure 6.1 from Price [4, p. 196], caption: "Tracks related to four known impact structures. Examples a, b and c are of *certain* impacts. (a) is that for the Popigai crater, diameter 100 km and age 34.6 Ma. (b) is that of Chicxulub, diameter about 200 km and age 66.25 Ma. (c) is that of Manicouagan, 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 $\epsilon \leq 1$ 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

$$V = \sqrt{\epsilon\,\frac{m}{M}}\,v \tag{14}$$

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

$$V_f^2 = V_i^2 + 2as \tag{15}$$

where $V_i$ is given by (14) and $V_f = 0$ when the plate stops moving. Solving for the distance $s$, we obtain

$$s = \frac{\epsilon}{2\,\mu_k g}\,\frac{m}{M}\,v^2\,. \tag{16}$$

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 $s = 420\,\mathrm{m}$ from (16). For $\epsilon = 0.5$ implying that 50% of the collision energy is available as kinetic energy, we get an initial plate velocity $V_i = 22\,\mathrm{m/s}$ from (14) and a lithospheric slip displacement $s = 210\,\mathrm{m}$ 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: $F_a > F_s = \mu_s Mg$.
- 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: $mv\cos\theta/\Delta t_p > \mu_s Mg$. 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 $s = \epsilon mv^2/2\mu_k gM$, 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 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).

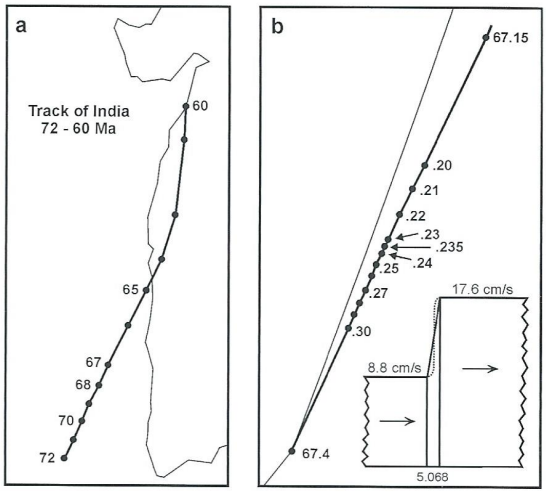



Fig. 3: Figure 6.7 from Price [4, p. 202], caption: "(a) Track of 'Bombay' over the period 72-60 Ma. It can be inferred from the distances between the points representing specific times that there was an abrupt change in velocity of the plate at about 67 Ma. (b) A detail of the track shown in (a) reveals 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 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;
- $\nabla p$ 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 $\rho$ 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; $\Delta h$ 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 $\alpha$ depended only on the time $t$):

$$\begin{aligned} m \, \frac{\Delta \vec{v}_{xy}}{\tau_0 + T_F} + \int_{t-\tau_0}^{t+T_F} D \, \frac{\vec{v}_{xy} \, \dot{V} \, d\tau}{\Delta V} = \\ = - \int_{t-\tau_0}^{t+T_F} \rho \, \nabla (g_{eff} \, \Delta h) \, \dot{V} \, d\tau \end{aligned} \quad (6)$$

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 $\tau_0$;
- $V(t)$ 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 $\omega$ the frequency of droplet self-mode oscillation and $\lambda$ 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 $\tau_0$ 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 $\tau_0$ 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 $\phi$ 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 $\rho$ is the density of the bath, $p$ is the acoustic pressure, $d_t \vec{v}$ is the convective temporal derivative of the moving fluid, $\beta$ 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:

$$m\,\Delta v_{xy} = H\,k \tag{10}$$

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 $\mathbf{u}$,

$$\boldsymbol{\xi} = \nabla \times \mathbf{u}, \tag{1}$$

also known as the vorticity $\boldsymbol{\xi}$. Because the rotation occurs in the $x$-$y$ plane, the vertical component $\zeta$ 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 $\Omega$. 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,

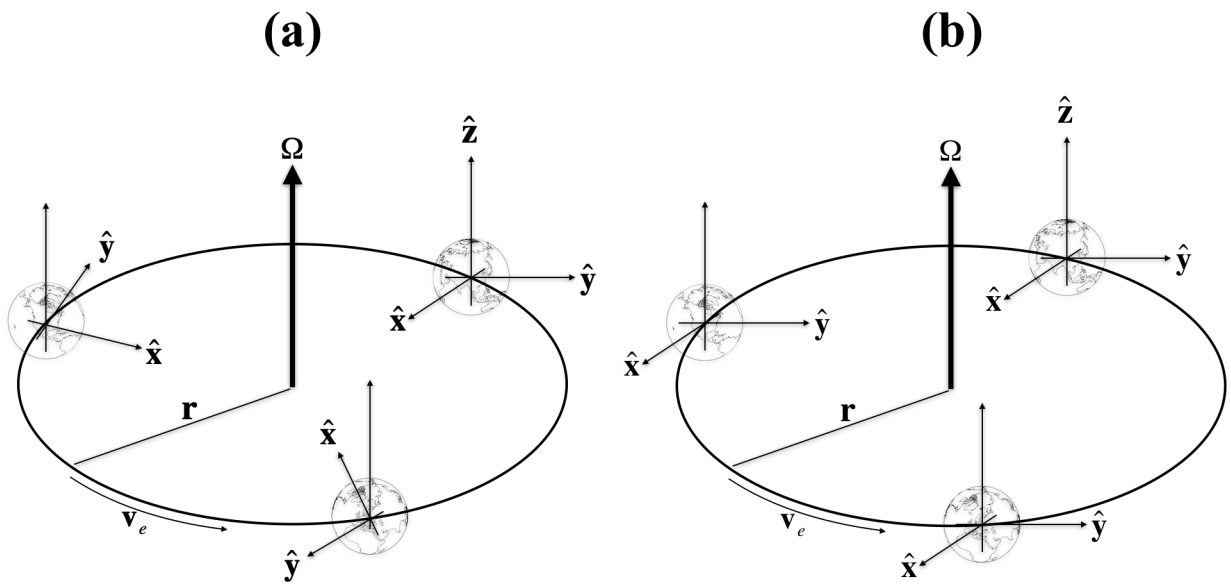



Fig. 1: **Coordinate systems.** An orbiting coordinate system around the stationary Sun is period locked (a).The unit vectors maintain a fixed orientation (b) by the clockwise annual spin of the celestial sphere about its axis $\hat{\mathbf{z}}$ to oppose the counterclockwise spin-orbit locking about $\Omega$. The origin of the celestial sphere moves at constant angular velocity about the axis $\Omega$ 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" effect at a distance, the frame-dependent Coriolis circulation, is what aberration of light may represent and could be an equivalent to Snell's law.

*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].

## 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

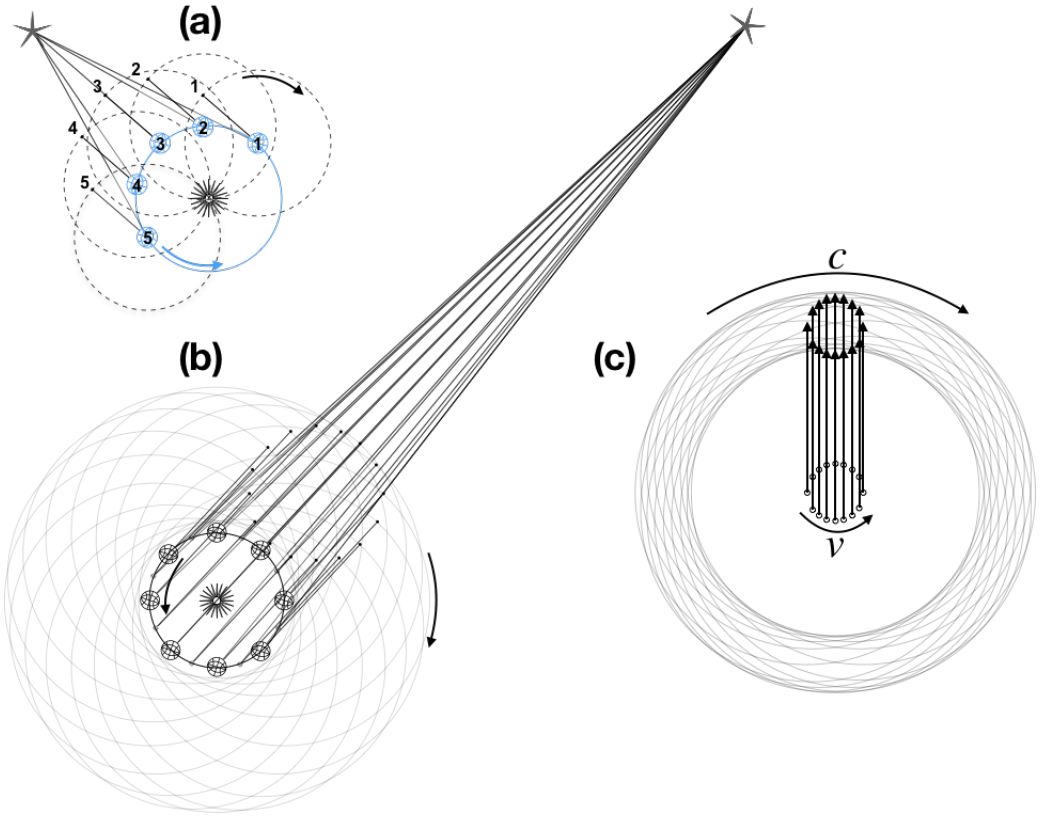



Fig. 2: **Celestial sphere radius and the velocity field.** Views of a planetary body rotating in a stationary frame of reference with the Sun as its centre. The celestial sphere centered on the planet with the same radius as the orbit (a) has a fixed orientation. The spindle of the celestial sphere 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

$$OQ = R\cos\theta + C\cos\phi, \tag{5}$$

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,

$$R\sin\theta = C\sin\phi, \tag{6}$$

corresponds to the tangential component of stellar aberration. Substitution of (6) into (5) provides the combined form, independent of the subtending angle $\phi$ term, where the angle $\theta$ equals the angular velocity $\dot{\theta}$ of the planet at time $t$, yielding

$$OQ = R\cos\theta + \sqrt{C^2 - R^2\sin^2\theta}. \tag{7}$$

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 $\phi$, when observing aberration in the direction of the plane of the ecliptic. The second term at the right hand side of (7), normalized to $C$,

$$\cos\phi = \sqrt{1 - \frac{R^2}{C^2}\sin^2\theta}, \tag{8}$$

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 $\phi$ (Eqs. 6, 8) to

$$\sin\phi = \beta\sin\theta \tag{9}$$

and

$$\cos\phi = \sqrt{1 - \beta^2\sin^2\theta}. \tag{10}$$

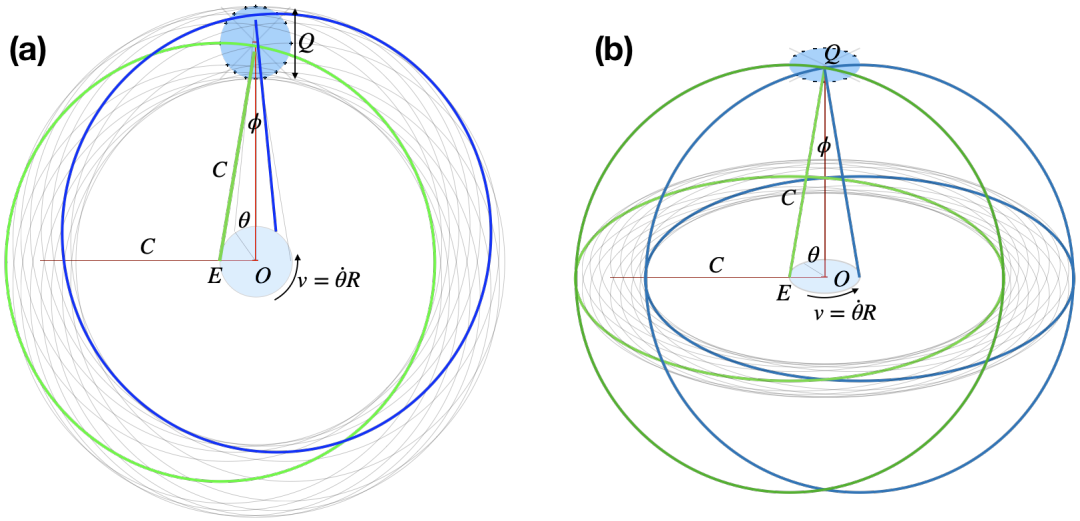



Fig. 3: **Geometry of negative parallax.** Point $O$ represents the Sun and $E$ is the position of the Earth in its orbit. Earth is the centre of the celestial sphere, the light blue circle is the Earth orbital plane and the dark blue filled circle is the traced-out orbital path of the Earth at distance $C$. (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 $\beta$ is a Lorentz triangle with sides $1, \beta$ and $\gamma\beta$, where

$$\begin{aligned} \sin\phi &= \beta, \\ \cos\phi &= \sqrt{1-\beta^2} = \gamma^{-1}. \end{aligned} \tag{11}$$

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 $\beta$ 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 $\beta$ 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 $\gamma$ and $\beta$ (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.

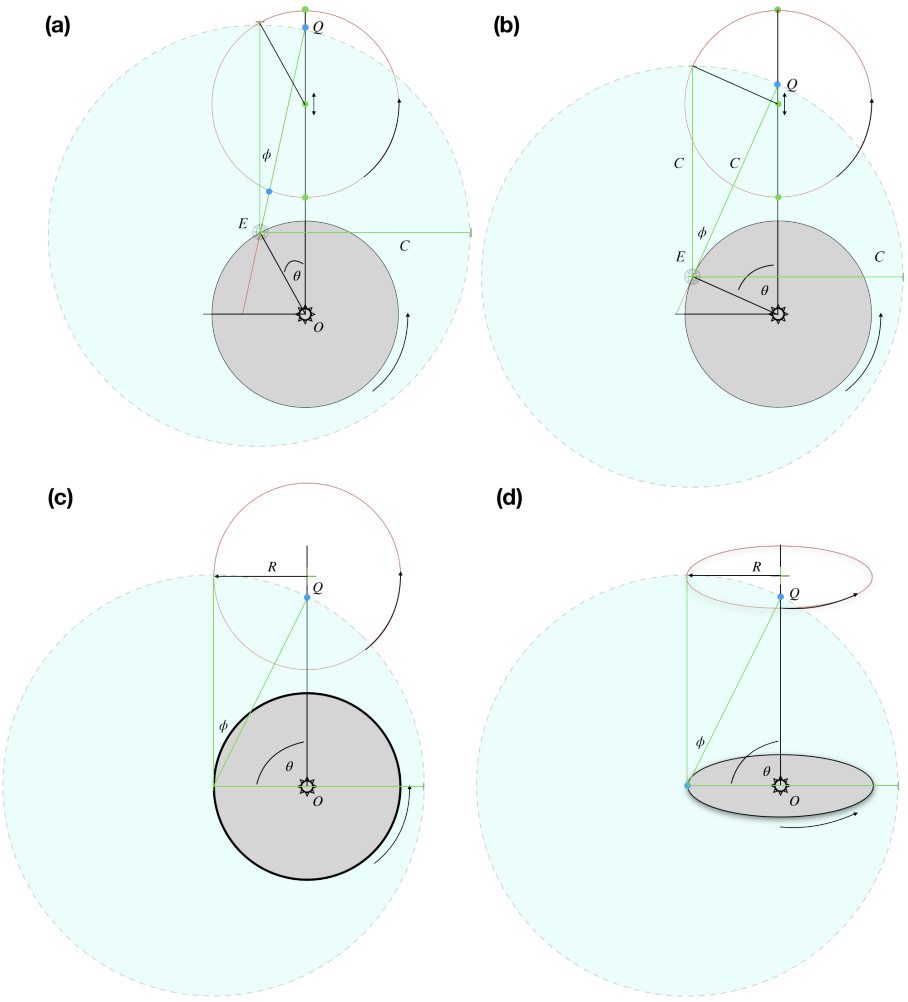



Fig. 4: **Scotch yoke reciprocating motion.** Earth's celestial sphere (light blue) with respect to the origin $O$, (the Sun) at (a), (b) and (c), affects the position of point $Q$, which is the imaginary Sun (see Section 2) as seen from Earth, $E$, sliding it up and down along the vertical line from $O$ 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 $\beta$ 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 $\beta$, 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 $\phi$ [arcsec]. The duration of a light signal from the celestial sphere surface to the planet $\mu^{-1}$ (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 $\phi_{bc}$ (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 |
| $\phi$ | 33.19 | 24.09 | 20.49 | 16.6 | 8.99 | 6.64 | 4.68 | 3.74 |
| $\mu^{-1}$ | 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 |
| $\phi_{bc}$ | 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 $\beta$, 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 ~58 days ago and from ~1 602 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 ($\beta^1$, cf. Fig. 2C), a radial field of view of 0.002 seconds of an arc ($\beta^2$), a radial field of view of 0.2 microseconds of an arc ($\beta^3$), 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 ~30.6 km/s [4], which resembles Earth's orbital velocity. The values reported are consistent

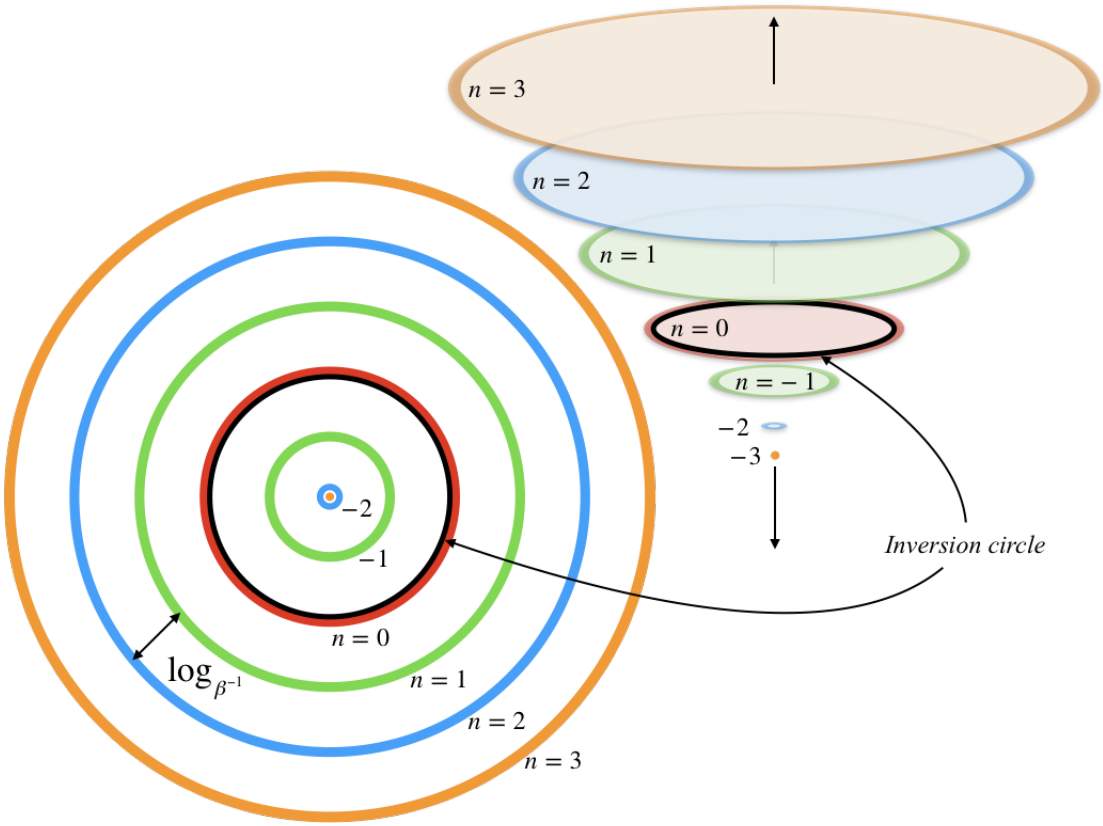



Fig. 5: **Inversion Circle.** The logarithmic scale with base $\beta^{-1}$ separates the circles (left and right panels). The inversion circle (black) is identical to the planetary orbit (red) when $n = 0$ ( 24). The green circles represent a circle with radius $R/\beta$ equal to $C$ when $n = 1$, the principal celestial sphere radius as described in this paper, and its inverse with the inverse radius $\beta R$ 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 $\beta^3$. 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 $\mu$arcsec 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 identical, 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

*en.wikipedia.org/wiki/Barycentric_coordinates_(astronomy)

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

$$C_n = \beta^{-n} R, \tag{21}$$

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))

$$OQ = R\cos\theta \pm R\cos\theta, \tag{22}$$

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

$$r_k = \beta^{2+n} C_n \tag{23}$$

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 $\Omega$ 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

$$\begin{aligned} u_C &= \Omega \times C_n = \Omega \times \beta^{-n} R = \beta^{-n} v \\ u'_C &= \Omega \times C'_n = \Omega \times \beta^{n} R = \beta^{n} v. \end{aligned} \tag{24}$$

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

$$r'_k = \beta^{-2} R \tag{25}$$

as defined by $C_2$. The tangential velocities of points and their inverse points (the coordinates) obey

$$u_C\, u'_C = v^2 \Leftrightarrow \frac{u'_C}{v} = \frac{v}{u_C}. \tag{26}$$

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°, $\beta$ 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 $1.57 \times 10^{15}$ 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

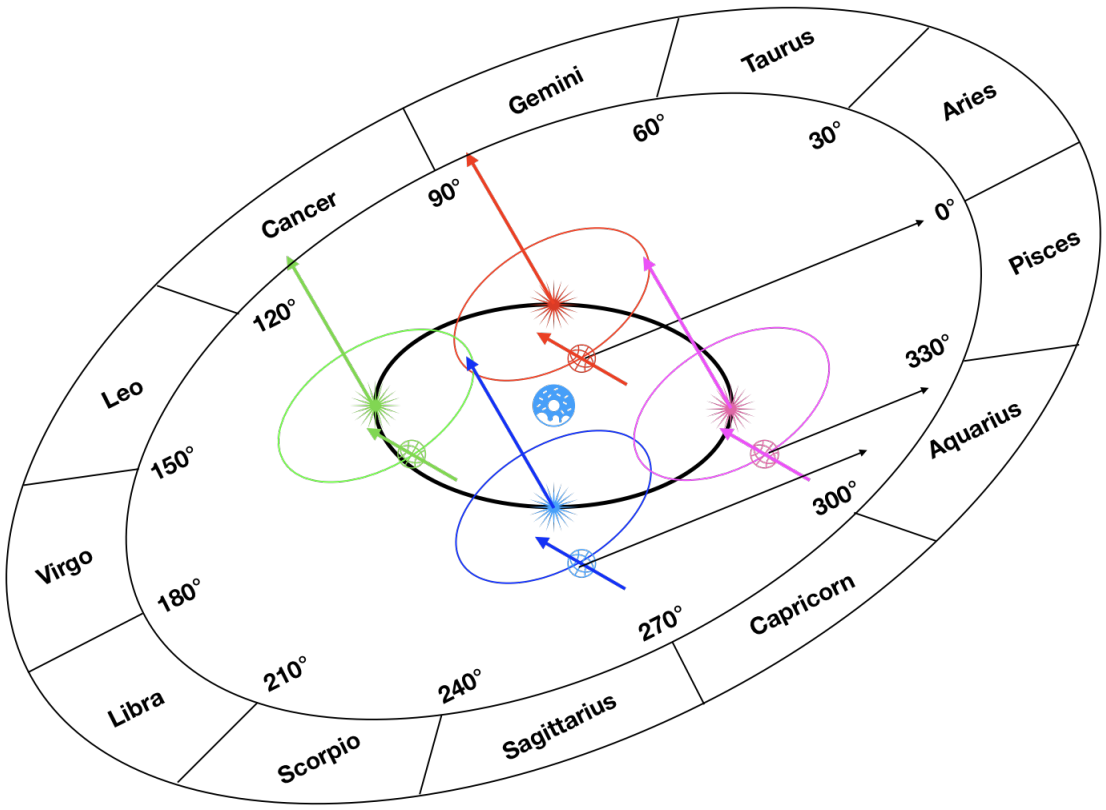



Fig. 6: **Stellar aberration by the precession of the equinoxes.** The orbit of Earth around the Sun is shown at four different positions of the Sun in its orbit about the equinoctial center shown as a blue-colored donut shaped symbol. The color coded arrows centered on Earth depicts the north celestial 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,

$$\omega t' = k\mathbf{r}', \tag{27}$$

where $\omega$ is the angular frequency and $k$ is the angular wavenumber of the waveform, to

$$\omega \sec\phi \left(t - \mathbf{r}c^{-1}\sin\phi\right) = k\sec\phi\,(\mathbf{r} - ct\sin\phi)\,. \tag{28}$$

Because $\omega k^{-1} = c$, the above identity after rearranging becomes

$$\omega t \sec\phi\,(1+\sin\phi) = k\mathbf{r}\sec\phi\,(1+\sin\phi)\,. \tag{29}$$

We note that the velocity $u$ of the wave is not affected because

$$u = \frac{\mathbf{r}}{t} = \frac{\omega\sec\phi\,(1+\sin\phi)}{k\sec\phi\,(1+\sin\phi)} = c, \tag{30}$$

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

$$z + 1 = \sec\phi(1+\sin\phi) = \sqrt{\frac{1+\beta\sin\theta}{1-\beta\sin\theta}}\,. \tag{31}$$

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 $\phi$. 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 $v/c$,

$$z + 1 = \frac{\omega_{receiver}}{\omega_{source}} = \frac{v}{c} + 1, \tag{32}$$

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, $\theta$ and $\beta$, 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)),

$$\frac{\omega_r^2}{\omega_s^2} = \frac{1 + \beta\sin\theta}{1 - \beta\sin\theta} \Rightarrow \beta\sin\theta = \frac{\omega_r^2 - \omega_s^2}{\omega_r^2 + \omega_s^2}\,. \tag{33}$$

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

$$\frac{u}{c} = \frac{c}{u'} \Rightarrow uu' = c^2, \tag{34}$$

where $u \leq c$ and $u' \geq c$. We get

$$\frac{\omega_r^2}{\omega_s^2} = \frac{u' + c\sin\theta}{u' - c\sin\theta} = \frac{c^2/u + c\sin\theta}{c^2/u - c\sin\theta} = \frac{c + u\sin\theta}{c - u\sin\theta}\,, \tag{35}$$

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 **L**, the Galilean **G** and the light retardation **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.

Received on July 4, 2020

## 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 $\Delta t_d$ for time change and two points for cause and effect, resulting in a delay of time $\Delta t_d$ 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 $\Delta t_d$. The existing quantum mechanical state with 4-momentum $p_\mu$ at a point $x^\alpha$ in space-time 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 $\Delta x^\alpha$ is a timelike 4-vector, meaning cause-effect delation in space-time. The time component of $\Delta x^\alpha$ is an amount representing the cause-effect time delation $\Delta t_d$ 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.

$$\text{past-future cause-effect vector}: x^\mu\phi_p(x^\alpha+\Delta x^\alpha),$$

$$\text{future-past cause-effect vector}: (x^\mu+\Delta x^\mu)\phi_p(x^\alpha).$$

### 2.2 Difference of cause-effect vectors

Since there are two states between $\Delta x^\alpha$ as discussed above, by combining them, the state change can be of two types:

$$\text{type 1}: (x^\mu+\Delta x^\mu)\phi_p(x^\alpha)-x^\mu\phi_{p'}(x^\alpha+\Delta x^\alpha)\ . \qquad (1)$$

$$\text{type 2}: x^\mu\phi_{p'}(x^\alpha+\Delta x'^\alpha)-(x^\mu-\Delta x^\mu)\phi_p(x^\alpha)\ . \qquad (2)$$

*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.

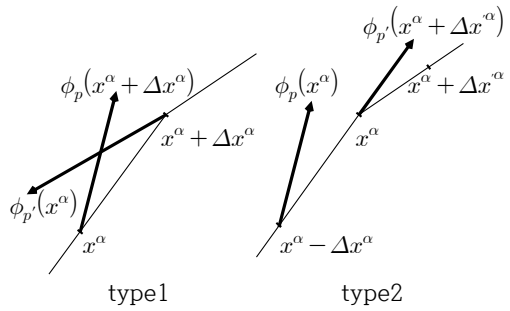



Fig. 1: Schematics of type 1 and type 2

## 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 $x^\alpha$ has the same magnitude of contribution at $x^\alpha - \Delta x^\alpha$ and $x^\alpha + \Delta x^\alpha$.

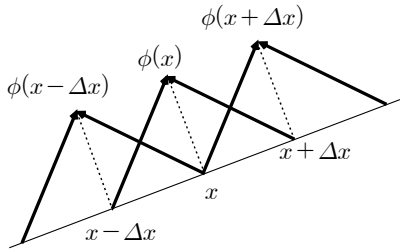



Fig. 2: Contributions to each $\phi$

In Fig. 2, $\phi(x)$ is a mixture of contributions from $x - \Delta x$ and $x + \Delta x$. The same applies to the other $\phi$'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 $x^\mu$ 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.

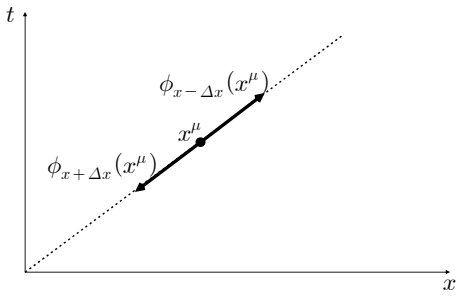



Fig. 3: two contributions to $\phi(x)$

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 : $\Psi$ and $\phi$ are analytic functions.

But in reality it is discontinuous and it is difficult to figure this out. This assumption approximates discontinuous $\Psi$ and $\phi$ as $C^\infty$ 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.

$$\begin{aligned}
&(x^\mu + \Delta x^\mu)\phi(x^\alpha) - x^\mu \phi(x^\alpha + \Delta x^\alpha) \\
&= (x^\mu + \Delta x^\mu)\phi(x^\alpha) - x^\mu \sum_{n=0}^{\infty} \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n \phi(x^\alpha) \\
&= \Delta x^\lambda \left\{\delta^\mu_\lambda \phi(x^\alpha) - x^\mu \frac{\partial \phi(x^\alpha)}{\partial x^\lambda}\right\} - \\
&\quad - x^\mu \sum_{n=2}^{\infty} \frac{1}{n!} \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n \phi(x^\alpha) \ .
\end{aligned}$$

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

$$\begin{aligned}
&(x^\mu + \Delta x^\mu)\,\phi(x^\alpha) - x^\mu \phi(x^\alpha + \Delta x^\alpha) \\
&= -\left[x^\mu, \Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right]\phi(x) - \\
&\quad - \left[x^\mu, \sum_{n=2}^{\infty} \frac{1}{n!}\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n\right]\phi(x) \\
&= -\left[x^\mu, \sum_{n=1}^{\infty} \frac{1}{n!}\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n\right]\phi(x) \\
&= -\left[x^\mu, \exp\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right) - 1\right]\phi(x) \\
&= -\left[x^\mu, \exp\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)\right]\phi(x) .
\end{aligned}$$

For the progress of the calculation, we define the 4-momentum operator $P_\lambda$, and commutation relation of $x^\mu$ and $P_\lambda$ as follows.

$$\begin{aligned}
P_\lambda &\equiv i\hbar \frac{\partial}{\partial x^\lambda} \\
[x^\mu, P_\lambda] &\equiv -i\hbar \delta^\mu_\lambda
\end{aligned} \tag{9}$$

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).

$$\begin{aligned}
[x_i, F(P_i)] &= i\hbar \frac{\mathrm{d}F}{\mathrm{d}P_i} \\
[x_0, F(P_0)] &= -i\hbar \frac{\mathrm{d}F}{\mathrm{d}P_0} .
\end{aligned} \tag{10}$$

$$\begin{aligned}
&(x^\mu + \Delta x^\mu)\,\phi(x^\alpha) - x^\mu \phi(x^\alpha + \Delta x^\alpha) \\
&= \Delta x^\mu \exp\left(-\frac{i}{\hbar}\Delta x^\alpha P_\alpha\right)\phi(x) .
\end{aligned} \tag{11}$$

Therefore, the equation for spinor function $\Psi_a(x)$ is

$$\begin{aligned}
&(x^\mu + \Delta x^\mu)\,\Psi_a(x) - x^\mu \Psi_a(x + \Delta x) \\
&= \Delta x^\mu \exp\left(-\frac{i}{\hbar}\Delta x^\alpha P_\alpha\right)\Psi_a(x) .
\end{aligned} \tag{12}$$

### 3.2 Type 2

After a similar calculation process as in type 1, the equation that corresponds to (12) is (16).

$$\begin{aligned}
&x^\mu \phi(x + \Delta x) - (x^\mu - \Delta x^\mu)\,\phi(x) \\
&= x^\mu \sum_{n=0}^{\infty} \frac{1}{n!}\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n \phi(x) - (x^\mu - \Delta x^\mu)\,\phi(x) \\
&= \Delta x^\alpha \left(x^\mu \frac{\partial}{\partial x^\alpha} + \delta^\mu_\alpha\right)\phi(x) + x^\mu \sum_{n=2}^{\infty} \frac{1}{n!}\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n \phi(x) .
\end{aligned}$$

For $n \geq 2$

$$x^\mu \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n = \left\{x^\mu, \left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n\right\}_{\text{anticommutation}} .$$

$$\begin{aligned}
&x^\mu \phi(x + \Delta x) - (x^\mu - \Delta x^\mu)\,\phi(x) \\
&= \left\{x^\mu, \Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right\}\phi(x) + \left\{x^\mu, \sum_{n=2}^{\infty} \frac{1}{n!}\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n\right\}\phi(x) \\
&= \left\{x^\mu, \sum_{n=1}^{\infty} \frac{1}{n!}\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n\right\}\phi(x) \\
&= \left\{x^\mu, \exp\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right) - 1\right\}\phi(x) .
\end{aligned} \tag{13}$$

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.

$$\begin{aligned}
P_\lambda &\equiv i\hbar \frac{\partial}{\partial x^\lambda} \\
\{x^\mu, P_\lambda\} &\equiv i\hbar \delta^\mu_\lambda .
\end{aligned} \tag{14}$$

And using the following (15), the final equation corresponding to (12) of type 1 is the following (16).

$$\begin{aligned}
\{x_i, G(P_i)\} &= -i\hbar \frac{\mathrm{d}G}{\mathrm{d}P_i} \\
\{x_0, G(P_0)\} &= i\hbar \frac{\mathrm{d}G}{\mathrm{d}P_0} .
\end{aligned} \tag{15}$$

$$\begin{aligned}
&x^\mu \phi(x + \Delta x) - (x^\mu - \Delta x^\mu)\,\phi(x) \\
&= \left(\Delta x^\mu \exp\left(-\frac{i}{\hbar}\Delta x^\alpha P_\alpha\right) - 2x^\mu\right)\phi(x) .
\end{aligned} \tag{16}$$

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.

$$\begin{aligned}
U(t, t_0) &= \exp\left(-\frac{iH(t - t_0)}{\hbar}\right) \\
i\hbar \frac{\partial U(t, t_0)}{\partial t} &= HU(t, t_0) \\
i\hbar \frac{\partial}{\partial t} U(t, t_0)\,|\,\alpha, t_0\,\rangle &= HU(t, t_0)\,|\,\alpha, t_0\,\rangle .
\end{aligned}$$

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

$$\begin{aligned} &(x^\mu + \Delta x^\mu)\Psi_a(x) - x^\mu \Psi_a(x + \Delta x) \\ &= \Delta x^\mu \exp\left(-\frac{i}{\hbar}\Delta x^\alpha \hat{P}_\alpha\right)\Psi_a(x) \\ &\equiv \Delta x^\mu U(\Delta x)\Psi_a(x) . \end{aligned} \tag{17}$$

Suppose $U(\Delta x)$ in (17) is an evolution operator by $\Delta x$ in space-time. Then

$$\begin{aligned} &\Psi_a^p(\Delta x) \equiv U(\Delta x)\Psi_a^p(0) \\ &= \exp\left(-\frac{i}{\hbar}\Delta x \cdot \hat{P}\right)\Psi_a^p(0) = \exp\left(-\frac{i}{\hbar}\Delta x \cdot \ p\right)\Psi_a^p(0) . \end{aligned} \tag{18}$$

Successive evolution by $n\Delta x^\alpha = x^\alpha$ can be expressed as below.

$$\begin{aligned} &\exp\left(-\frac{i}{\hbar}\Delta x \cdot \hat{P}\right) \cdot \cdots \cdot \exp\left(-\frac{i}{\hbar}\Delta x \cdot \hat{P}\right)\Psi_a^p(0) \\ &= \exp\left(-\frac{i}{\hbar}n\Delta x \cdot p\right)\Psi_a^p(0) = \exp\left(-\frac{i}{\hbar}x \cdot p\right)\Psi_a^p(0) \\ &= \exp\left(-\frac{i}{\hbar}x \cdot \hat{P}\right)\Psi_a^p(0) = U(x)\Psi_a^p(0) . \end{aligned}$$

Therefore, $U(x)$ satisfies the following equation.

$$\Psi_a^p(x) = U(x)\Psi_a^p(0) . \tag{19}$$

If we apply the Klein-Gordon operator to (19) and use $p$ as constant, we get the following result ($\hbar = 1$).

$$\begin{aligned} &\left(\partial_\mu \partial^\mu + m^2\right)\Psi_a^p(x) \\ &= \left(\partial_\mu \partial^\mu + m^2\right)e^{-ix\cdot\hat{P}}\Psi_a^p(0) \\ &= \left(\partial_\mu \partial^\mu + m^2\right)e^{-ix\cdot p}\Psi_a^p(0) \\ &= \left(-p_\mu p^\mu + m^2\right)e^{-ix\cdot p}\Psi_a^p(0) \\ &= 0 . \end{aligned} \tag{20}$$

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

$$\begin{aligned} &x^\mu \phi(x + \Delta x) - (x^\mu - \Delta x^\mu)\phi(x) \\ &= \left(\Delta x^\mu exp\left(-\frac{i}{\hbar}\Delta x^\alpha P_\alpha\right) - 2x^\mu\right)\phi(x) \\ &\equiv \Delta x^\mu V \phi(x) . \end{aligned} \tag{21}$$

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

$$\Delta x^\mu V \simeq -2x^\mu \tag{22}$$

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

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = 0 \tag{23}$$

where $\tau$ 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.

$$\begin{aligned} &x^\mu\left(\phi(x) + \Delta x^\alpha \frac{\partial \phi}{\partial x^\alpha}\right) - (x^\mu - \Delta x^\mu)\phi(x) \\ &= \Delta x^\mu \left\{\phi(x) - \frac{i}{\hbar}\Delta x^\alpha p_\alpha \phi(x)\right\} - 2x^\mu \phi(x) \\ &\Delta x^\alpha \frac{\partial \phi(x)}{\partial x^\alpha} = -2\phi(x) . \end{aligned}$$

$$\therefore \phi(x) \propto \exp\left(-2\frac{\Delta x \cdot x}{\Delta x \cdot \Delta x}\right) . \tag{24}$$

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 $\phi_p(x) = V(x)\,\phi_p(0)$.

$$\begin{aligned} &\left(\partial_\mu \partial^\mu + m^2 f\right)\left(e^{-ix\cdot\hat{P}} - 2\right)\phi_p(0) = 0 \,. \\ &\therefore f(x) = \frac{e^{-ix\cdot p}}{e^{-ix\cdot p} - 2} \,. \end{aligned} \tag{27}$$

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.

$$\begin{aligned} &\left(\partial_\mu \partial^\mu + m^2 f(x)\right)\phi(x) = 0 \\ &\left(\partial_\mu \partial^\mu + m^2 f^*(x)\right)\phi^*(x) = 0 \,. \end{aligned} \tag{28}$$

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.

$$\begin{aligned} &\phi(x) \to e^{-iq\theta(x)}\phi(x) \,,\ \phi^*(-x) \to e^{iq\theta(-x)}\phi^*(-x) \,. \\ &\frac{\delta\phi(x)}{\delta\theta(x)} = -iq\phi(x) \,,\ \frac{\delta\phi^*(-x)}{\delta\theta(x)} = iq\phi^*(-x)\,\frac{\delta\theta(-x)}{\delta\theta(x)} \,. \end{aligned} \tag{31}$$

According to Noether's theorem, if the action is invariant under gauge transformations, there is a vanishing divergence current, whose value is

$$\begin{aligned} J^\mu &= -\left[\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi(x)\right)}\frac{\delta\phi(x)}{\delta\theta(x)} + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\phi^*(-x)\right)}\frac{\delta\phi^*(-x)}{\delta\theta(x)}\right] \\ &= iq\left[\phi(x)\,\partial^\mu\phi^*(-x) - \phi^*(-x)\,\partial^\mu\phi(x)\,\frac{\delta\theta(-x)}{\delta\theta(x)}\right] \,. \end{aligned} \tag{32}$$

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 $t = 0, \vec{\mathrm{p}} = (p, 0, 0)$ 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 $\epsilon$, 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

$$\text{simple pole} \quad z_0 = -\frac{(\epsilon + ip)}{\epsilon^2 + p^2} \ln 2 .$$

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.

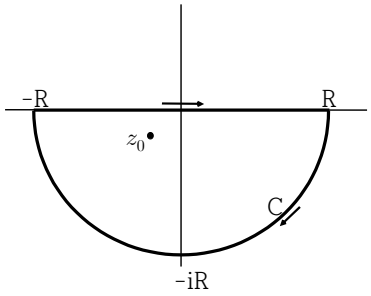



Fig. 4: Contour of integration

$$\begin{aligned} &\oint dz \frac{e^{i(p+i\epsilon)|z|}}{e^{i(p+i\epsilon)z} - 2} \\ &= \lim_{R\to\infty} \int_{-R}^{R} dx \frac{e^{i(p+i\epsilon)|x|}}{e^{i(p+i\epsilon)x} - 2} + \int_C dz \frac{e^{i(p+i\epsilon)|z|}}{e^{i(p+i\epsilon)z} - 2} \\ &= 2\pi i a_{-1} = 2\pi i \cdot 2^i . \end{aligned} \tag{37}$$

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.

$$T^{00} = \left( \vec{\mathrm{p}}^2 + m^2 \right) \phi^* \phi .$$

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.

$$\begin{aligned} &\mathcal{L} = \partial_\mu \phi^* (-x) \partial^\mu \phi (x) - m^2 f (x) \phi^* (-x) \phi (x) \\ &T^{\mu\lambda} = \partial^\mu \phi^* (-x) \partial^\lambda \phi (x) + \partial^\mu \phi (x) \partial^\lambda \phi^* (-x) - \eta^{\mu\lambda} \mathcal{L} . \end{aligned} \tag{40}$$

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.

Received on July 24, 2020

## References


1. Whitrow G. J. The Natural Philosopy 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].

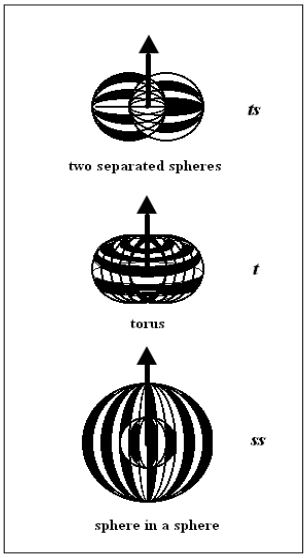



Fig. 1: Observables structures of the electron. The arrow represents its magnetic moment

## 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

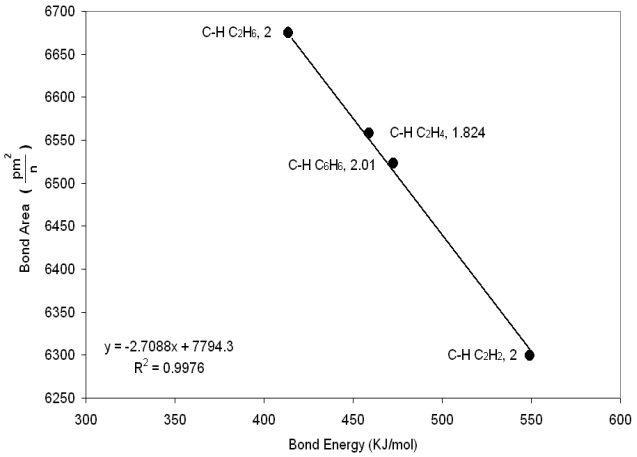



Fig. 2: Bond area vs. bond energy for C-H bonds in different molecules.

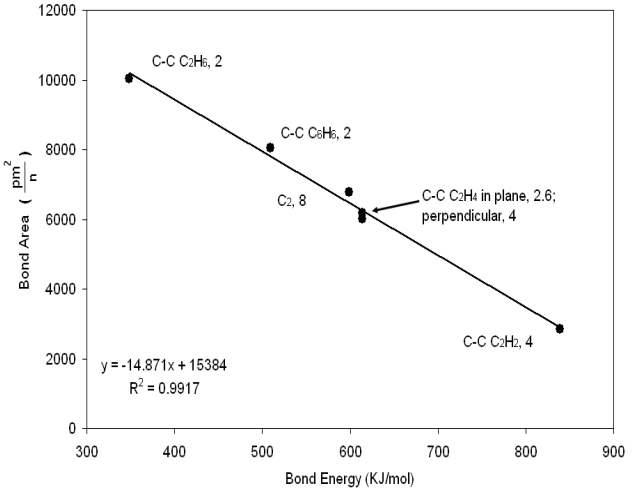



Fig. 3: Bond area vs. bond energy for C-C bonds in different molecules.

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-

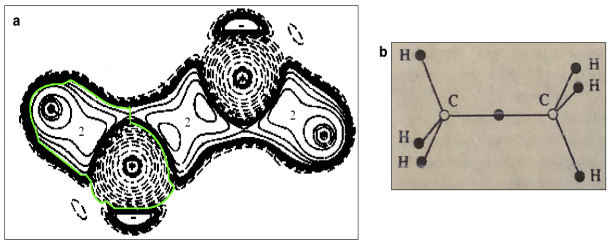



Fig. 4: **a**) Contour map of the Ethane molecule and its electron count. The green line shows how the C-H bond was cut. The C-C bond was also cut accordingly. Reprinted and adapted with permission from [6]. Copyright (1996) American Chemical Society. **b**) Silvi's et al. structure from [2]. It presents 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 $\sigma_{C-H}$ - $\sigma_{C-H}*$ 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 $\mu_0$ is the permeability of the free space, $m_e$ is the electron magnetic moment, $r$ is the distance between magnetic moments and $\theta$ 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.

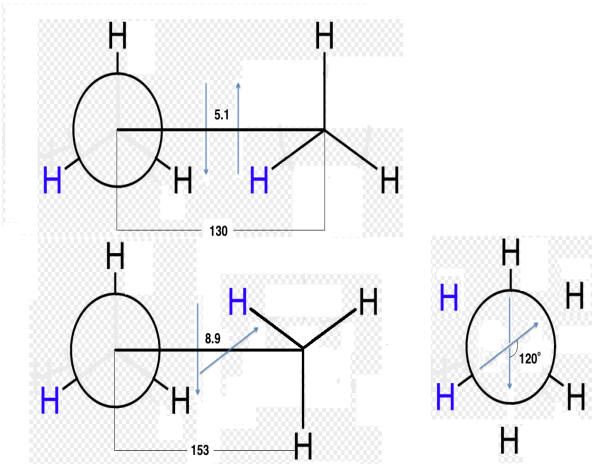
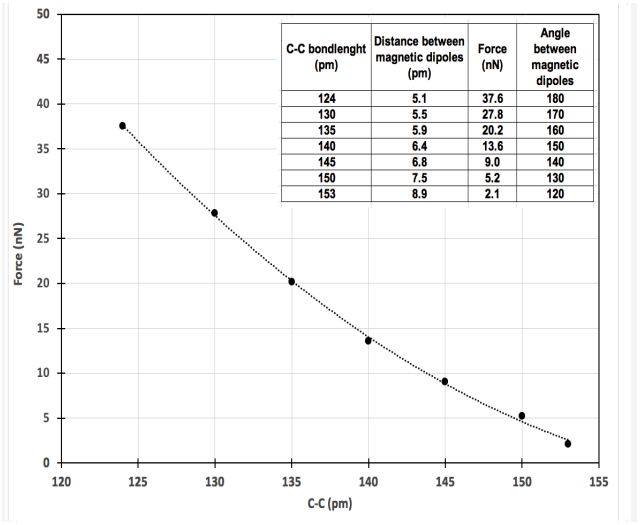



Fig. 5: Eclipsed (top) and Staggered (down) ethane conformers. The eclipsed conformer present the maximum magnetic attraction and shorter C-C bond length. Whereas, the staggered conformer has the lowest magnetic attraction and the longest C-C bond length. The extra projection shown down right is to present 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.

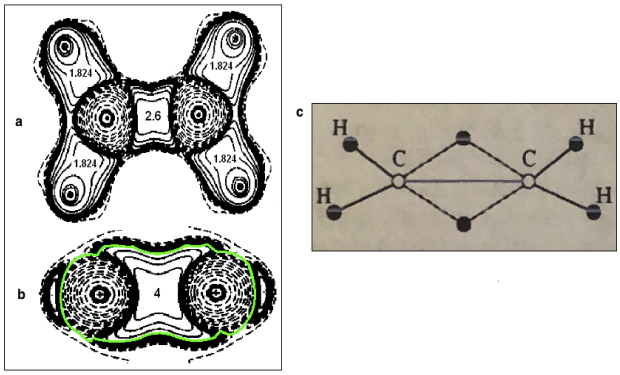



Fig. 7: Contour map of the Ethene molecule and its electron count. The C-C bond electron count at the plane of the nuclei **(a)** is different from the count at the perpendicular plane **(b)**. The green line shows how the C-C bond was cut. Reprinted (adapted) with permission from [6]. Copyright (1996) American Chemical Society. **c**) Silvi's et al. structure from [2] used with permission of the publisher.

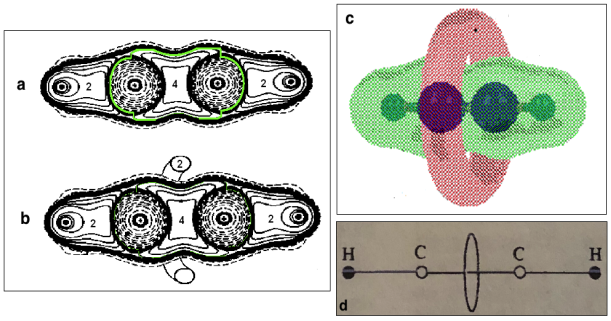



Fig. 8: **a**) Contour map of the ethyne molecule and its electron count, there is a lack of two electrons. The green line shows how the C-C bond was cut. **b**) These two electrons are fused in a toroidal lone pair around the C-C bond. Reprinted (adapted) from [6]. Copyright (1996) American Chemical 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

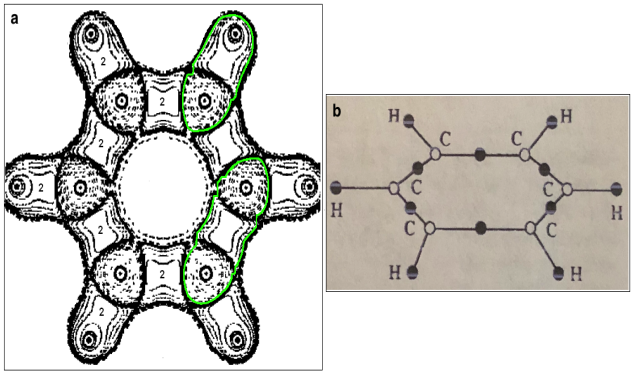



Fig. 9: **a**) Contour map of the Benzene molecule and its electron count. There is a lack of six electrons. These electrons are fused in two lone pairs at both sides of the $C_6$ ring. The green line shows how the bonds were cut. This is from [7] used under Creative Commons license. **b**) Silvi's et 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 $\pi$ 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.

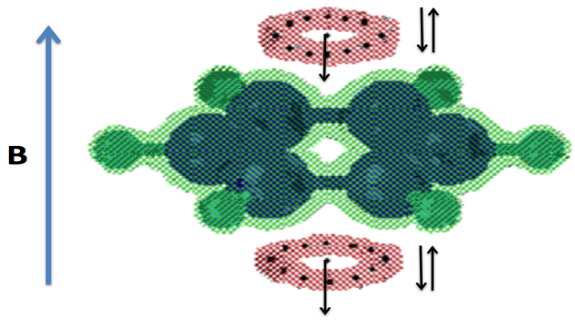



Fig. 10: Molecular electrostatic potential of benzene. The arrows depict the coupling of the three electrons in each lone pair. When an external magnetic field **B** is imposed, the magnetic moments of the two odd electrons aligned against it as shown.This is from [9] used with permission of the publisher.

## 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

$$g = C_1\left(\frac{\alpha}{\pi}\right) + C_2\left(\frac{\alpha}{\pi}\right)^2 + C_3\left(\frac{\alpha}{\pi}\right)^3 + C_4\left(\frac{\alpha}{\pi}\right)^4 + C_5\left(\frac{\alpha}{\pi}\right)^5 ... \tag{3}$$

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.

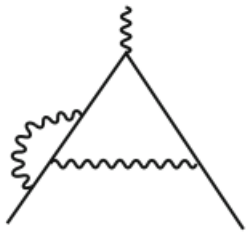


Fig. 1: Feynman diagram IIc.

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).

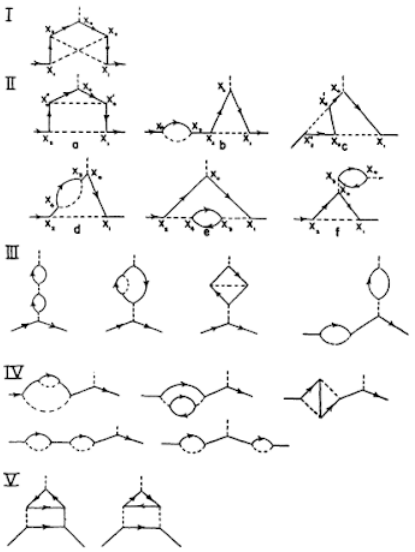



Fig. 2: Feynman diagrams.

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

$$C_2 = \mathrm{I} + \mathrm{IIa} + \mathrm{IIc} + \mathrm{IId} + \mathrm{IIe} = -2,973\,. \qquad (6)$$

| 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 | $\pi^2$ | $\pi^2 \ln 2$ | $\zeta(3)$ | 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.

| 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.

## 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 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. | $\pi^2$ | $\pi^2$ 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

$$C_2 = \frac{197}{144} + \frac{1}{12}\pi^2 - \frac{1}{2}\pi^2 ln(2) + \frac{3}{4}\zeta(3) = -0.328\,. \qquad (10)$$

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]

$$II_c = -\frac{323}{24} + \frac{31}{9}\pi^2 - \frac{49}{6}\pi^2 ln(2) + \frac{107}{4}\zeta(3) \qquad (11)$$

[Petermann]

$$II_c = -\frac{67}{24} + \frac{1}{18}\pi^2 + \frac{1}{3}\pi^2 ln(2) - \frac{1}{2}\zeta(3) \qquad (12)$$

The calculation of each of the four factors in diagram IIc is shown in the following table:

| | Const. | $\pi^2$ | $\pi^2$ $ln(2)$ | $\zeta(3)$ | 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

$\Lambda$ – 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:

$$\rho = \frac{3c^2}{4\pi G a^2} \tag{4}$$

$$p = -\frac{1}{3}c^2\rho\,. \tag{5}$$

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 $a\,\mathrm{d}\psi$ ($0 \le \psi \le \pi$). We get it by integration:

$$V = a^3\,4\pi \int_0^\pi \sin^2\psi\,\mathrm{d}\psi = 2\pi^2 a^3\,. \tag{6}$$

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

$$a_0 = \sqrt{\frac{G\hbar}{3\pi c^3}} \cong 5.26 \times 10^{-36}\,\mathrm{m}\,. \tag{9}$$

The minimum time interval then:

$$t_0 = \frac{a_0}{c} = \sqrt{\frac{\hbar G}{3\pi c^5}} \cong 1.76 \times 10^{-44}\,\mathrm{s}\,. \tag{10}$$

The first quantum mass $M_0$ of the universe is given by the relation:

$$M_0 = \sqrt{\frac{3\pi\hbar c}{4G}} \cong 3.34 \times 10^{-8}\,\mathrm{kg}\,. \tag{11}$$

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:

$$M_{\bullet_0} = \sqrt{\frac{\hbar c}{4G}} \cong 1.09 \times 10^{-8}\,\mathrm{kg}\,. \tag{13}$$

Thus, the initial quantum was below the event horizon, which began at a distance given by the minimum size of the black hole:

$$a_{\bullet_0} = \sqrt{\frac{G\hbar}{c^3}} \cong 1.62 \times 10^{-35}\,\mathrm{m}\,. \tag{14}$$

A black hole of this mass is characterized by temperature:

$$T_{\bullet_0} = \frac{\hbar c^3}{8\pi k G M_{\bullet_0}} = 1.13 \times 10^{31}\,\mathrm{K}\,. \tag{15}$$

## 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:

$$w = v + \imath\dot{a} \tag{24}$$

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 $\alpha$ 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 $\alpha_v$ in the relation (26) is the standard form of Einstein coefficient $\alpha$. 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:

$$\rho_i = \frac{3c^2}{4\pi G\rho(\frac{a}{\dot{a}})^2} = \frac{3\dot{a}^2}{8\pi G a^2} = \frac{3H^2}{8\pi G} = \rho_k \tag{29}$$

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:

$$\mathrm{d}U = -p\mathrm{d}V = c^2\mathrm{d}M\,. \tag{31}$$

Mass movement in the direction of the expansion of the universe and its rise with time induce a force, which has size:

$$F = \imath^2\frac{\mathrm{d}M}{\mathrm{d}t}c = -\frac{3\pi c^4}{2G}\,. \tag{32}$$

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 $T_r = 2.726\,\mathrm{K}$ results the energy density $U \cong 4.18\times10^{-14}\,\mathrm{J\,m^3} \cong 0.26\,\mathrm{eV\,cm^3}$

out of the relation (34). This corresponds to the measured value of the density of the relict radiation $0.25\,\mathrm{eV\,cm^3}$.

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 $T_r = 2.726\,\mathrm{K}$ 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:

$$a_i = \sqrt{2} n^2 a_0 \cong 1.27 \times 10^{26}\,\mathrm{m}\,. \tag{38}$$

The Hubble constant $H$ is then according to (29):

$$H = \frac{c}{a} \cong 2.35 \times 10^{-18}\,\mathrm{s^{-1}} \cong 72.63\,\mathrm{km\,s^{-1}\,Mpc^{-1}}\,. \tag{39}$$

This value is consistent with the value of the Hubble constant determined in 2018: $H = 73.52 \pm 1.62\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The actual age of the universe is therefore:

$$t = \frac{1}{H} \cong 13.5 \times 10^9\,\mathrm{yrs}\,. \tag{40}$$

The mass of the universe (from the point of view of an internal observer) is:

$$M_i = n^2 M_0 \cong 5.72 \times 10^{53}\,\mathrm{kg}\,. \tag{41}$$

## 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)

$$v^2 = \frac{G M_g}{r} \tag{43}$$

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.

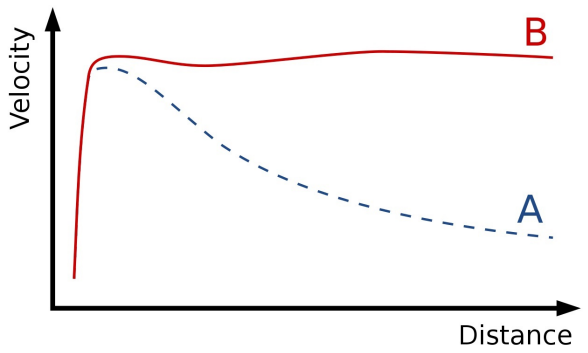



Fig. 1: Dependence of orbital velocity on distance from center of galaxy

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:

$$M_{0n} = \frac{n^2 M_0}{n^3} = \frac{M_0}{n} \cong 8.08 \times 10^{-39}\,\mathrm{kg}\,. \tag{49}$$

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:

$$\begin{aligned} \Delta a_t &= \sqrt{(\Delta a_0)^2 + \left(\frac{\Delta(m_0 \dot{a})}{m}\, t\right)^2} \\ &= \sqrt{a_0^2 + \left(\frac{m_0 c t_0}{m_0}\right)^2} = a_0\sqrt{2}\,. \end{aligned} \tag{51}$$

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 $\Delta t_r$:

$$\frac{\Delta t_r}{\Delta t_{e-today}} = \frac{\lambda_r}{\lambda_{e-today}} = 1 + z\,. \tag{58}$$

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:

$$r_e = c \sin\left(\int_{t_e}^{t_r} \frac{\mathrm{d}t}{a}\right) = \sin\left(\ln \frac{t_r}{t_e}\right) = \sin\left[\ln(1+z)\right]. \tag{62}$$

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 $L \cong 2.765 \times 10^{28}\,\mathrm{W}$. It acknowledges that the model above can correspond with our reality.

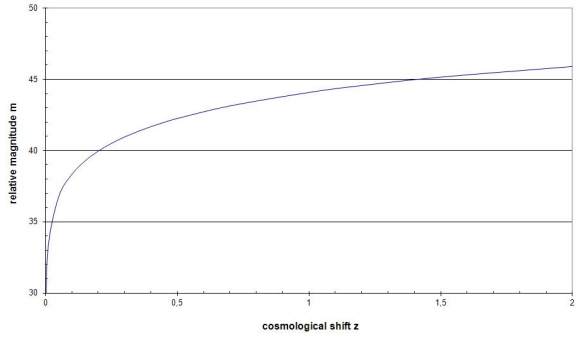



Fig. 2: Relative supernova magnitude – calculated for $L = 2.765 \times 10^{28}\,\mathrm{W}$

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).

$$\Delta(m - M) = 5 \log\left(\frac{r_e}{r_{e0}}\right). \tag{64}$$

$$r_{e0} = \sinh\left[\ln(1+z)\right]. \tag{65}$$

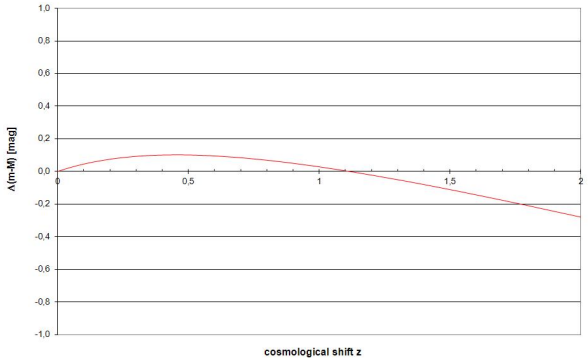



Fig. 3: Residual Hubble diagram – without consideration of dust influence

## 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.

Received on July 22, 2020

## 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 $\mathrm{SL}(2,\mathbb{R})$ with the subgroup of rotation $\mathrm{SO}(1,1)^{+}$. 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 $\mathrm{SL}(4,\mathbb{R})$ 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 $\phi^{-1}$ 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 $\eta$ 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

$$J_{\mu a} = \frac{\partial x^a}{\partial \overline{x}^\mu} \quad J^{\mu a} = \frac{\partial \overline{x}^a}{\partial x^\mu}\,. \tag{1}$$

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 $\eta$ has the signature $(-\,+\,+\,+)$. The transformation is non-singular $\mathrm{J} \neq 0$, the matrix is bijective, and the inverse transform represents the simple inverse matrix $\overline{\mathrm{J}} = \mathrm{J}^{-1}$. If the order of indexes as per (1), the equation can be written in the matrix notation (for both covariant and contravariant forms) as

$$\mathrm{g} = \mathrm{J} \cdot \eta \cdot \mathrm{J}^{\mathrm{T}}\,. \tag{3}$$

The capital letters are used for matrices excluding the metric tensor g, and Minkowski $\eta$. In matrix notation, the form (covariant or contravariant) will be specified in the text. For the spherical symmetry case, the Jacobian matrices are $4 \times 4$

$$\begin{bmatrix} \mathrm{J} & 0 \\ 0 & \mathrm{I}_2 \end{bmatrix}.$$

Therefore, J can be written as $2 \times 2$ 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

$$v = \mathrm{th}(\beta) = \sqrt{\frac{r_g}{r}}$$
$$\sinh(\beta) = \frac{v}{\sqrt{1-v^2}}$$
$$\gamma = \cosh(\beta) = \frac{1}{\sqrt{1-v^2}} = \frac{1}{\sqrt{1-\frac{r_g}{r}}}$$

## 3 Step one: from Minkowski to Gulfstrand-Painlevé

The first coordinate transformation as given in [5, 22] is

$$dx^1 = d\overline{x}^1 - v d\overline{x}^0 \qquad dx^0 = d\overline{x}^0 \tag{4}$$

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

$$\begin{aligned} J^{(1)}_{\mu a} &= \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{th}(\beta) \\ 0 & 1 \end{bmatrix} \\ J^{\mu a}_{(1)} &= \begin{bmatrix} 1 & 0 \\ v & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \mathrm{th}(\beta) & 0 \end{bmatrix} \end{aligned} \tag{5}$$

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 $v = \mathrm{th}(\beta)$ and the (imaginary) shear angle is $\beta$ 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

$$\begin{aligned} \mathrm{g}_1 = \mathrm{J}^{(1)} \cdot \eta \cdot \mathrm{J}^{(1)\mathrm{T}} &= \begin{bmatrix} -\left(1 - v^2\right) & v \\ v & 1 \end{bmatrix} \\ &= \begin{bmatrix} -\left(1 - \frac{r_g}{r}\right) & \sqrt{\frac{r_g}{r}} \\ \sqrt{\frac{r_g}{r}} & 1 \end{bmatrix}. \end{aligned} \tag{6}$$

## 4 Step two: to the Schwarzschild metric

The second coordinate transformation $\mathrm{J}_2$ is pull-back from the comoving G-P frame to the coordinate frame of reference redefining time coordinate. The covariant form is

$$\mathrm{J}^{(2)} = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \tag{7}$$

where $b$ is the *arbitrary* parameter †. The total coordinate transformation is the product of both transforms

$$\mathrm{J} = \mathrm{J}^{(2)} \cdot \mathrm{J}^{(1)} = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & v \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & v \\ b & vb + 1 \end{bmatrix} \tag{8}$$

that leads to the metric tensor

$$\mathrm{g} = \mathrm{J} \cdot \eta \cdot \mathrm{J}^{\mathrm{T}} = \begin{bmatrix} -(1 - v^2) & (v^2 - 1)\, b + v \\ (v^2 - 1)\, b + v & (vb + 1)^2 - b^2 \end{bmatrix}. \tag{9}$$

Choosing $b$ in the way to eliminate the off-diagonal terms one obtains the Schwarzschild metric

$$\begin{aligned} g_{\mu\nu} &= \begin{bmatrix} -\left(1 - v^2\right) & 0 \\ 0 & \left(1 - v^2\right)^{-1} \end{bmatrix} \\ &= \begin{bmatrix} -\cosh^{-2}(\beta) & 0 \\ 0 & \cosh(\beta) \end{bmatrix}. \end{aligned} \tag{10}$$

After $b$ has been defined, the second transformation becomes

$$\begin{aligned} \mathrm{J}^{(2)}_{\mu \mathrm{a}} &= \begin{bmatrix} 1 & 0 \\ \sinh(\beta)\cosh(\beta) & 1 \end{bmatrix} \\ \mathrm{J}^{\mu \mathrm{a}}_{(2)} &= \begin{bmatrix} 1 & \sinh(\beta)\cosh(\beta) \\ 0 & 1 \end{bmatrix}. \end{aligned} \tag{11}$$

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

$$J_{\mu a} = \frac{\partial x^a}{\partial \overline{x}^{\mu}} = \begin{bmatrix} \frac{\partial x_0}{\partial \overline{x}_0} & \frac{1}{i}\frac{\partial x_1}{\partial \overline{x}_0} \\ i\,\frac{\partial x_0}{\partial \overline{x}_1} & \frac{\partial x_1}{\partial \overline{x}_1} \end{bmatrix}. \tag{12}$$

In such a way, the Jacobian matrices constitute the subgroup of $\mathrm{SL}(2,\mathbb{C})$ with only two imaginary off-diagonal elements in the matrices. Let's denote this group as $\mathrm{SL}(2,\mathbb{C})^* \in \mathrm{SL}(2,\mathbb{C})$. Then, considering only the real parts, there is one-to-one mapping of $Z' \in \mathrm{SL}(2,\mathbb{C})^*$ to $Z \in \mathrm{SL}(2,\mathbb{R})$ as follows

$$Z' = \begin{bmatrix} a & -i\,b \\ i\,c & d \end{bmatrix} \to Z = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{13}$$

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 $\mathrm{SL}(2,\mathbb{R})$. Introduced in such a way, the group $\mathrm{SL}(2,\mathbb{C})^*$ is isomorphic to $\mathrm{SL}(2,\mathbb{R})$. This mapping is multiplicative and a bijection. Hence, all operations in $\mathrm{SL}(2,\mathbb{R})$ can be translated to $\mathrm{SL}(2,\mathbb{C})^*$ and vice versa using this isomorphism. Such mapping allows one to utilize $\mathrm{SL}(2,\mathbb{R})$ on the Lorentz/Minkowski basis $\mathbb{H}^{1(2)}$, instead of its default, the Euclidean basis $\mathbb{R}^2$.

*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".

‡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 $K \in SO(2)$.

One may see that using the mapping (13), the elements of these three groups become the matrices of the following form

$$N = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} \quad A = \begin{bmatrix} k & 0 \\ 0 & k^{-1} \end{bmatrix} \quad K = \begin{bmatrix} \cosh(\alpha) & \sinh(\alpha) \\ \sinh(\alpha) & \cosh(\alpha) \end{bmatrix} \tag{14}$$

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 $\in$ N (upper triangular matrices). Therefore, the decomposition can be applied to contravariant $J_{(1)}$ and to contarvariant $J^{(2)}$ 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

$$J^{\mu a}_{(1)} = \begin{bmatrix} 1 & 0 \\ \mathrm{th}(\beta) & 0 \end{bmatrix} = \begin{bmatrix} 1 & -\sinh(\beta)\cosh(\beta) \\ 0 & 1 \end{bmatrix} \cdot \cdot \begin{bmatrix} \cosh(\beta) & 0 \\ 0 & \cosh^{-1}(\beta) \end{bmatrix} \cdot \begin{bmatrix} \cosh(\beta) & \sinh(\beta) \\ \sinh(\beta) & \cosh(\beta) \end{bmatrix}. \tag{15}$$

Notably, that N in the factorization becomes already known matrix $N = J^{-1}_{(2)}$ (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 $K \equiv \Lambda$ 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*

$$A^{\nu b} = \begin{bmatrix} \cosh(\beta) & 0 \\ 0 & \cosh^{-1}(\beta) \end{bmatrix}$$

$$A_{\nu b} = \begin{bmatrix} \cosh^{-1}(\beta) & 0 \\ 0 & \cosh(\beta) \end{bmatrix}. \tag{18}$$

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‡

$$J^{\nu a}_{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ v_x & 1 & 0 & 0 \\ v_y & 0 & 1 & 0 \\ v_z & 0 & 0 & 1 \end{bmatrix}. \tag{19}$$

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 $\theta$ and $\phi$ 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 ($\eta$ 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 $|\mathrm{J}| = 1$ 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 $|\mathrm{J}| = 1$ 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 $\mathrm{SL}(2, \mathbb{R})$ 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.

$\mathrm{SL}(2, \mathbb{R})$ 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".

**Acknowledgements**

I would like to express profound gratitude to professor Keith Conrad for the valuable comments and the critical notes. I am also thankful to professor Andrew Hamilton for useful discussion on the approach.

Submitted on Sept. 9, 2020

## 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.Ẇhat 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.Ṗ.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 $\beta$ 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 ($\nu$) 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]

$$-\hbar^2 c^2 \frac{\partial^2 \overline{\psi}(q)}{\partial q^2} + \left(m^2 c^4 - E^2\right) \overline{\psi}(q) = 0 \tag{2}$$

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 $\nu_e$ and progressively greater for the analogous charged current interactions for $\nu_\mu$ and $\nu_\tau$. 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 $\{\overline{\psi}, \overline{\check{\psi}}\} = \{\exp(+ikq), \exp(-ikq)\}$ 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

$$\partial_q \overline{\psi} \,/\, \overline{\psi} \;=\; \partial_q \ln(\overline{\psi}) \;=\; ik \tag{3}$$

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]

$$\begin{aligned}
\overline{\psi}_2 &= \overbrace{\underbrace{\alpha \exp(+ikq)}_{\text{latent incident wave}} + \underbrace{\beta \exp(-ikq)}_{\text{latent reflected wave}}}^{\text{bispectral solution of SKGE by superpositional principle}} && (4)\\
&= \underbrace{(\alpha - \beta)\exp(ikq)}_{\text{traveling wave}} + \underbrace{2\beta \cos(kq)}_{\text{standing wave}} \\
&= \underbrace{(\alpha + \beta)\cos(kq) + i(\alpha - \beta)\sin(kq)}_{\text{coherent standing waves}} && (5)\\
&= \underbrace{A_{\overline{\psi}} \exp(i\, P_{\overline{\psi}})}_{\text{compoundly modulated traveling wave}}, && (6)
\end{aligned}$$

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

$$A_{\overline{\psi}} = \overbrace{[\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)]^{1/2}}^{\text{amplitude modulation}}$$

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 $\alpha$ and $\beta$ 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 $\alpha$ and $\beta$ are normalized by

$$\alpha^2 - \beta^2 = 1 \tag{7}$$

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

$$A_{\overline{\psi}}\big|_{\beta=0} = [\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)]^{1/2}|_{\beta=0} = \alpha|_{\beta=0} = 1 \tag{8}$$

and

$$P_{\overline{\psi}}\big|_{\beta=0} = \arctan\left(\frac{\alpha-\beta}{\alpha+\beta}\tan(kq)\right)\Big|_{\beta=0} = kq. \tag{9}$$

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 $2\alpha$ given by [27]

$$A_{\overline{\psi}}\big|_{\beta=\alpha} = 2\alpha\left(\frac{1+\cos(2kq)}{2}\right)^{1/2} = 2\alpha\cos(kq) \tag{10}$$

and

$$A_{\overline{\psi}}\big|_{-\beta=\alpha} = 2\alpha\left(\frac{1-\cos(2kq)}{2}\right)^{1/2} = 2\alpha\sin(kq) \tag{11}$$

consistent with (5). The corresponding phase would be

$$P_{\overline{\psi}}\big|_{\beta=\alpha} = \arctan\left(\frac{\alpha-\beta}{\alpha+\beta}\tan(kq)\right)\Big|_{\beta=\alpha} = 0 \tag{12}$$

and

$$P_{\overline{\psi}}|_{-\beta=\alpha} = \arctan\left(\frac{\alpha-\beta}{\alpha+\beta}\tan(kq)\right)\Big|_{-\beta=\alpha} = \frac{\pi}{2} \tag{13}$$

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 $\beta$. 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 $\{\overline{\psi}, \breve{\overline{\psi}}\} = \{\exp(+ikq), \exp(-ikq)\} = \{\overline{\psi}, \psi\}$ 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 $\{\alpha(q), \beta(q)\}_{q_a \le q \le q_b}$ that are explicitly still subject to the normalization

$$\alpha^2(q) - \beta^2(q) = 1, \quad q_a \le q \le q_b,$$

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

$$\begin{aligned} \overline{\psi}_2(q_a) &= \alpha(q_a)\exp(ikq_a) + \beta(q_a)\exp(-ikq_a) \\ &= [1 + \beta^2(q_a)]^{1/2}\exp(kq_a) + \beta(q_a)\exp(-ikq_a) \\ &\xrightarrow{q \to q_b,\ \therefore \beta(q) \to 0} \exp(ikq_b), \quad q_a \le q \le q_b \end{aligned} \tag{14}$$

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

$$\beta(q) = \frac{\beta(q_a)}{2}\left[1 + \cos\left(\frac{q - q_a}{q_b - q_a}\pi\right)\right], \quad q_a \le q \le q_b.$$

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

$$T = \frac{\alpha^2 - \beta^2}{\alpha^2} = \frac{1}{\alpha^2} \quad \text{and} \quad R = \frac{\beta^2}{\alpha^2}, \tag{15}$$

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

$$\begin{aligned} &[1 + \beta^2(q_a)]^{1/2}\exp(kq_a) + \beta(q_a)\exp(-ikq_a) \\ &\qquad \xrightarrow{q \to q_b,\ \therefore \beta(q) \to \beta_b} \\ &(1 + \beta_b^2)^{1/2}\exp(ikq_b) + \beta_b\exp(-ikq_b) \end{aligned} \tag{16}$$

where $q_a < q \le q_b$. 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

$$\begin{aligned} A_{\overline{\psi}} &= \underbrace{[\alpha^2 + \beta^2 - 2\alpha\beta\cos(\pi - 2kq)]^{1/2}}_{\text{law of cosines}} \\ &= \underbrace{[\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)]^{1/2}}_{\text{law of cosines for exterior angles}}. \end{aligned}$$

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, $q \le q_a$, 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

$$Pm_{\overline{\psi}} = \arctan\left(\frac{\alpha - \beta}{\alpha + \beta}\tan(kq)\right) - kq, \quad q \le q_a \tag{18}$$

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]

$$\begin{aligned} \frac{\partial \arctan\left(\frac{\alpha-\beta}{\alpha+\beta}\tan(kq)\right)}{\partial q} &= \frac{(\alpha^2 - \beta^2)k}{\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)} \\ &= \frac{k}{\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)}. \end{aligned} \tag{19}$$

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

$$\begin{aligned} Am_{\overline{\psi}} &\equiv \frac{[\alpha^2 + \beta^2 + 2\alpha\beta\cos(2kq)]^{1/2} - (\alpha^2 + \beta^2)^{1/2}}{(\alpha^2 + \beta^2)^{1/2}} \\ &= \left[1 + \frac{2\alpha\beta\cos(2kq)}{\alpha^2 + \beta^2}\right]^{1/2} - 1, \quad q \le q_a. \end{aligned} \tag{20}$$

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

$$\begin{aligned} \partial_q\overline{\psi}_2 &= ik[\alpha\exp(ikq) - \beta\exp(-ikq)] \\ &= k[(\alpha - \beta)\cos(kq) - i(\alpha + \beta)\sin(kq)]\exp(i\pi/2) \\ &= k\underbrace{[\alpha^2 + \beta^2 - 2\alpha\beta\cos(2kq)]^{1/2}}_{\text{law of cosines}} \\ &\quad \times \exp\left[i\arctan\left(\frac{\alpha + \beta}{\alpha - \beta}\tan(kq)\right) + i\frac{\pi}{2}\right]. \end{aligned} \tag{21}$$

A difference between (4)–(6) for $\overline{\psi}_2$ and (21) for $\partial_q\overline{\psi}_2$ is the change of the sign of $\beta$ and the phase shift $\pi/2$. A finite $\beta$ 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

$$\begin{aligned}\partial_q\overline{\psi}_2(kq) &= ik[\alpha\exp(ikq)-\beta\exp(-ikq)]\\ &= k\{\alpha\exp[i(kq+\pi/2)]+\beta\exp[-i(kq+\pi/2)]\}\\ &= k\,\overline{\psi}_2(kq+\pi/2).\end{aligned}\tag{23}$$

Eq. (23) can be generalized to

$$\begin{aligned}\partial_q\overline{\psi}_2(kq) &= k\overline{\psi}_2(kq+n_1\pi),\\ n_1 &= \pm1/2,\ \pm3/2,\ \pm5/2,\ \cdots.\end{aligned}\tag{24}$$

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

$$A_{\overline{\psi}'} = k[\alpha^2+\beta^2-2\alpha\beta\cos(2kq)]^{1/2},\quad q\le q_a\tag{25}$$

and

$$P_{\overline{\psi}'} = \arctan\left(\frac{\alpha+\beta}{\alpha-\beta}\tan(kq)\right)+\frac{\pi}{2},\quad q\le q_a.\tag{26}$$

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

$$Am_{\overline{\psi}'} = k\left[1-\frac{2\alpha\beta\cos(2kq)}{\alpha^2+\beta^2}\right]^{1/2}-k,\quad q\le q_a\tag{27}$$

and

$$Pm_{\overline{\psi}'} = \arctan\left(\frac{\alpha+\beta}{\alpha-\beta}\tan(kq)\right)-kq,\quad q\le q_a.\tag{28}$$

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 $\pi/k$. The momentum of $\overline{\psi}_2$, using (4), (7) and (21), is given by

$$\begin{aligned}p &= \frac{\int_0^{\pi/k}\overline{\psi}_2^{\dagger}(q)\,\frac{\hbar}{i}\partial_q\overline{\psi}_2(q)\,dq}{\int_0^{\pi/k}\overline{\psi}_2^{\dagger}(q)\overline{\psi}_2(q)\,dq}\\ &= \hbar\frac{k\int_0^{\pi/k}[\alpha^2-\beta^2+2\alpha\beta\sin(2kq)]\,dq}{\int_0^{\pi/k}[\alpha^2+\beta^2+2\alpha\beta\cos(2kq)]\,dq}\\ &= \hbar\frac{(\alpha^2-\beta^2)\pi}{(\alpha^2+\beta^2)\pi/k} = \frac{\hbar k}{\alpha^2+\beta^2}.\end{aligned}\tag{29}$$

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

$$[\alpha^2-\beta^2+2\alpha\beta\sin(2kq)]_{q=3\pi/(4k)} = \overbrace{\alpha^2-\beta^2}^{=1}-2\alpha\beta.\tag{30}$$

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

$$2\alpha\beta = 1.\tag{31}$$

The particular values of $\alpha$ 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

$$\{\alpha_{\text{threshold}},\beta_{\text{threshold}}\} = \left\{\left(\tfrac{2^{1/2}+1}{2}\right)^{1/2},\left(\tfrac{2^{1/2}-1}{2}\right)^{1/2}\right\}.\tag{32}$$

The logic relationship

$$\alpha\ <\ /\ >\ \alpha_{\text{threshold}}\iff |\beta|\ <\ /\ >\ |\beta_{\text{threshold}}|$$

between $\alpha$ and $\beta$ 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

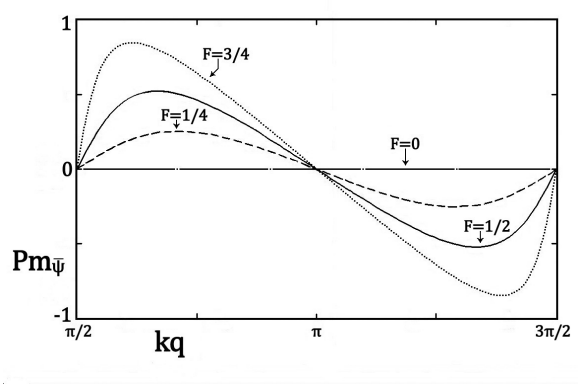



Fig. 1: The phase displacement due to phase modulation $Pm_{\overline{\psi}}$ as a function of $kq$ over a Riemann sheet for selected values of $F$. Both $Pm_{\overline{\psi}}$ and $kq$ are exhibited in units of radians.

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 $F$s 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 $\pi/2 \le kq \le 3\pi/2$ 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 $F$s 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

$$Pm_{\overline{\psi}}(\pi - kq) = -Pm_{\overline{\psi}}(\pi + kq),\ \ 0 < q < \pi/2.$$

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

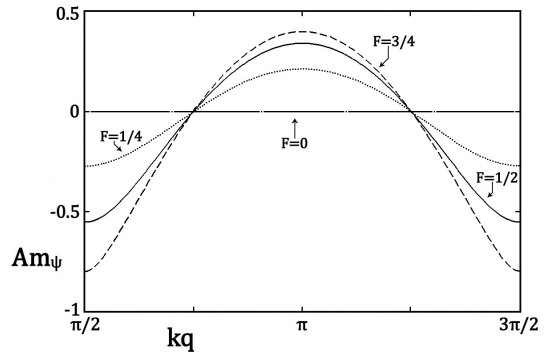



Fig. 2: The relative amplitude dilation $Am_{\overline{\psi}}$ as a function of $kq$ over a Riemann sheet for selected values of $F$. $Am_{\overline{\psi}}$ is dimensionless, and $kq$ is exhibited in units of radians.

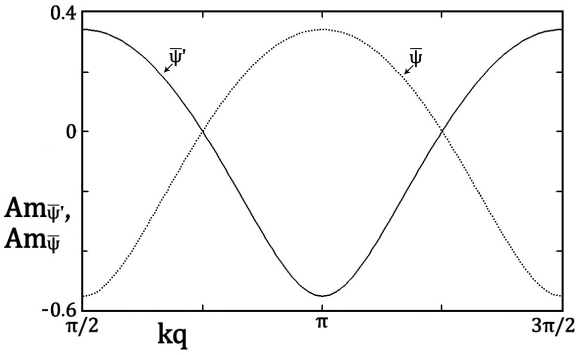



Fig. 3: The relative amplitude dilations due $Am_{\overline{\psi}}$ and $Am_{\overline{\psi}'}$ as functions of $kq$ over a Riemann sheet for $F = 1/2$. For an unbiased $Am_{\overline{\psi}'}$, $k = 1$ to facilitate comparison to dimensionless $Am_{\overline{\psi}}$. The amplitude modulations are dimensionless, 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

$$\left|\partial_q\overline{\psi}_2(kq)\right| : \left|\overline{\psi}_2(kq)\right| \rightsquigarrow \underbrace{\frac{\left|\partial_q\overline{\psi}_2(kq)\right|}{\left|\overline{\psi}_2(kq)\right|} = \frac{A_{\overline{\psi}'}(kq)}{A_{\overline{\psi}}(kq)}}_{\text{fractional form}}) = k\overbrace{\left(\frac{\alpha^2+\beta^2-2\alpha\beta\cos(2kq)}{\alpha^2+\beta^2+2\alpha\beta\cos(2kq)}\right)^{1/2}}. \tag{35}$$

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

$$\frac{A_{\overline{\psi}'}(kq)}{A_{\overline{\psi}}(kq)}\Big|_{\text{maximum}} = k\frac{\alpha+\beta}{\alpha-\beta} \text{ at } kq = \frac{\pi}{2},\ \frac{3\pi}{2}$$

and

$$\frac{A_{\overline{\psi}'}(kq)}{A_{\overline{\psi}}(kq)}\Big|_{\text{minimum}} = k\frac{\alpha-\beta}{\alpha+\beta} \text{ at } kq = \pi.$$

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 $\beta$ 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

$$\begin{aligned}\triangle P_{\overline{\psi}',\overline{\psi}}(kq) &= P_{\overline{\psi}'}(kq) - P_{\overline{\psi}}(kq)\\ &= P_{\overline{\psi}}(kq+\pi/2) + \pi/2 - P_{\overline{\psi}}(kq).\end{aligned} \tag{36}$$

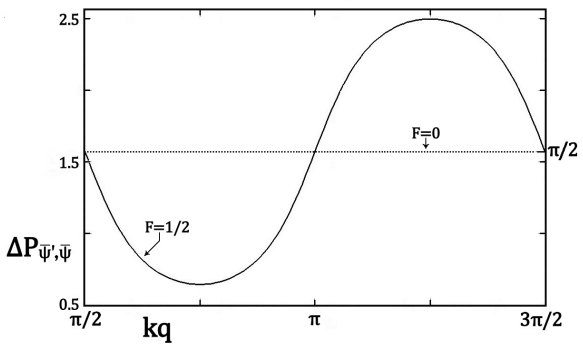



Fig. 4: The Phase difference $\triangle P_{\overline{\psi}',\overline{\psi}}(kq)$ as a function of $kq$ over a Riemann sheet for $F = 0, 1/2$. Both $\triangle P_{\overline{\psi}',\overline{\psi}}(kq)$ and $kq$ are exhibited in units of radians.

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 $kq = \pi/2,\ \pi,\ 3\pi/2$. 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,

$$(2n-1)\pi/2 \le kq \le (2n+1)\pi/2,\ n = 0, \pm1, \pm2, \cdots .$$

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 $2\pi$. Periodicity of $Am_{\overline{\psi}}$ is also given by

$$Am_{\overline{\psi}}(kq) = Am_{\overline{\psi}}(kq+n\pi),\ n = \pm1, \pm2, \pm3, \cdots .$$

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 $\{\alpha, \beta\} \to \{1, 0\}$ 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 $\beta$ 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 $\beta$ 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**

First, I heartily thank Marco Matone for his incisive critique of an earlier version of this paper. Second, this opus was self-funded by the author, who is unaffiliated.

Received on Sept. 22, 2020

## 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.

Progress in Physics is an American scientific journal on advanced studies in physics, registered with the Library of Congress (DC, USA): ISSN 1555-5534 (print version) and ISSN 1555-5615 (online version). The journal is peer reviewed and listed in the abstracting and indexing coverage of: Mathematical Reviews of the AMS (USA), DOAJ of Lund University (Sweden), Scientific Commons of the University of St.Gallen (Switzerland), Open-J-Gate (India), Referential Journal of VINITI (Russia), etc. Progress in Physics is an open-access journal published and distributed in accordance with the Budapest Open Initiative: this means that the electronic copies of both full-size version of the journal and the individual papers published therein will always be accessed for reading, download, and copying for any user free of charge.
The journal is issued quarterly (four volumes per year).

Electronic version of this journal: http://www.ptep-online.com



Postal address:
Department of Mathematics and Science, University of New Mexico,
705 Gurley Avenue, Gallup, NM 87301, USA