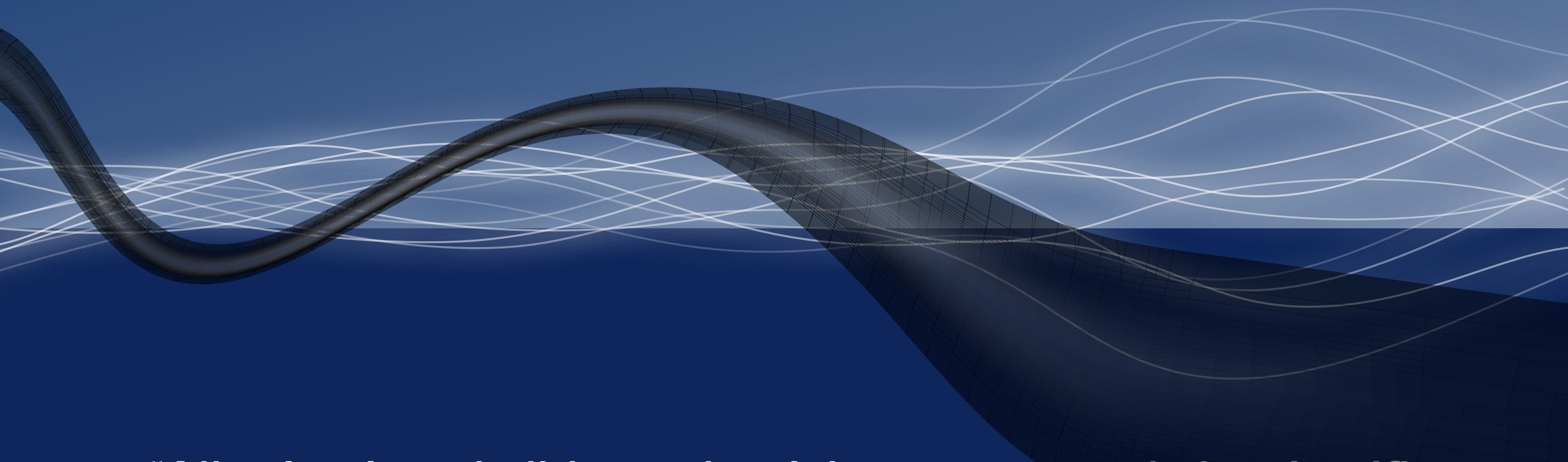
Issues 1-2 2021 | Volume 17

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

Advisory Board

Dmitri Rabounski,
Editor-in-Chief, Founder
Florentin Smarandache,
Associate Editor, Founder
Larissa Borissova,
Associate Editor, Founder

Editorial Board

Pierre Millette
millette@ptep-online.com
Andreas Ries
ries@ptep-online.com
Gunn Quznetsov
gunn@mail.ru
Ebenezer Chifu
chifu@ptep-online.com

Postal Address

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

Copyright © *Progress in Physics*, 2021

All rights reserved. The authors of the articles do hereby grant *Progress in Physics* non-exclusive, worldwide, royalty-free license to publish and distribute the articles in accordance with the Budapest Open Initiative: this means that electronic copying, distribution and printing of both full-size version of the journal and the individual papers published therein for non-commercial, academic or individual use can be made by any user without permission or charge. The authors of the articles published in *Progress in Physics* retain their rights to use this journal as a whole or any part of it in any other publications and in any way they see fit. Any part of *Progress in Physics* howsoever used in other publications must include an appropriate citation of this journal.

This journal is powered by LaTeX

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

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

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

April 2021 **Vol. 17, Issue 1**

CONTENTS

Information for Authors

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

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

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

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

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

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

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

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

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

Relativistically Correct Electromagnetic Energy Flow

Oliver Davis Johns
San Francisco State University, Department of Physics and Astronomy, Thornton Hall, 1600 Holloway Avenue, San Francisco, CA 94132 USA.
Email: ojohns@metacosmos.org

Detailed study of the energy and momentum carried by the electromagnetic field can be a source of clues to possible new physics underlying the Maxwell equations. But such study has been impeded by expressions for the parameters of the electromagnetic energy flow that are inconsistent with the transformation rules of special relativity. This paper begins by correcting a basic parameter, the local velocity of electromagnetic energy flow. This correction is derived by the direct application of the transformation rules of special relativity. After this correction, the electromagnetic energy-momentum tensor can then be expressed in a reference system comoving with the energy flow. This tensor can be made diagonal in the comoving system, and brought into a canonical form depending only on the energy density and one other parameter. The corrected energy flow and its energy-momentum tensor are illustrated by a simple example using static electric and magnetic fields. The proposal that electromagnetic momentum results from the motion of a relativistic mass contained in the fields is examined in the context of the corrected flow velocity. It is found that electromagnetic field momentum, though real, cannot be explained as due only to the motion of relativistic mass. The paper concludes that introducing the requirement of consistency with special relativity opens the study of electromagnetic energy and momentum to new possibilities.

1 Introduction

The Feynman example of a rotating disk with a magnet at its center and charged spheres on its perimeter provides a convincing argument that, to preserve the principle of angular momentum conservation, the field momentum of even a static electromagnetic field must be considered physically real.[*]

Since the energy density and momentum density of the electromagnetic field are real, it is important to investigate the details of the energy flow that they represent. Since special relativity is the symmetry theory of electrodynamics, it is essential that such investigations respect the transformation laws of special relativity.

In Section 2 a previously proposed candidate expression for a basic parameter, the velocity of energy flow at a given point in the electromagnetic field, is shown to be inconsistent with the transformation rules of special relativity and therefore incorrect. A corrected velocity expression is derived by explicit use of these rules.

Section 3 derives the electromagnetic energy-momentum tensor in a reference system comoving with this corrected velocity, and shows that it can be made diagonal and reduced to a canonical form that depends on two parameters derived from the values of the electric and magnetic fields.

Section 4 illustrates the results of the previous sections with an example using static electric and magnetic fields.

Section 5 considers the question of a relativistic mass density derived from the energy density of the electromagnetic field. It is found that this mass density does not correctly relate the momentum density to the flow velocity. Electromagnetic field momentum, although real, is not due only to the motion of relativistic mass.

Section 6 concludes that introducing the requirement of special relativistic covariance into the study of the flow of energy in electromagnetic fields opens up new possibilities for investigation of such flows.

Electromagnetic formulas in this paper are taken from Griffiths [6] and Jackson [7] with translation to Heaviside-Lorentz units. I denote four-vectors as $\mathsf{A} = A^0\mathsf{e}_0 + \mathbf{A}$ where e_0 is the time unit vector and the three-vector part is understood to be $\mathbf{A} = A^1\mathsf{e}_1 + A^2\mathsf{e}_2 + A^3\mathsf{e}_3$. In the Einstein summation convention, Greek indices range from 0 to 3, Roman indices from 1 to 3. The Minkowski metric tensor is $\eta_{\alpha\beta} = \eta^{\alpha\beta} = \mathrm{diag}(-1, +1, +1, +1)$. Three-vectors are written with bold type $\mathbf{A}$, and their magnitudes as A. Thus $|\mathbf{A}| = A$.

2 Velocity of energy flow

We begin with a basic parameter of the electromagnetic field. The flow velocity of the energy contained in the field at a given event can be defined as the velocity of a comoving observer who measures a zero energy flux there. Expressed in the precise language of Lorentz boosts:[†]

> The laboratory system coordinate velocity of the flow of electromagnetic field energy at a given event is the velocity $\mathbf{V}$ of a Lorentz boost that transforms the laboratory reference system into a reference system in which the Poynting energy

---

[*] Feynman et al [4], Section 17-4, Section 27-6, and Figure 17-5. Quantitative matches of field to mechanical angular momentum are found, for example, in Romer [12] and Boos [2].

[†] The Lorentz boost formalism used here is defined in Appendix A.

> flux vector is null at that event. An observer at that event and at rest in this system, which we call the comoving system and denote by primes, therefore measures a zero energy flux. The zero flux measurement indicates that this observer is comoving with the flow of energy. Such an observer has coordinate velocity $\mathbf{V}$ relative to the laboratory[*], and therefore $\mathbf{V}$ is the laboratory system coordinate velocity of the energy flow at the given event.

A previously proposed[†] candidate for the laboratory system coordinate velocity of the electromagnetic energy flow is the momentum density divided by the relativistic mass (defined as energy density divided by the square of the speed of light). Denoting this velocity as $\mathbf{u}_\mathrm{e}$ gives

$$\frac{\mathbf{u}_\mathrm{e}}{c} = \frac{\mathbf{G}}{(\mathcal{E}/c)} = \frac{2\,(\mathbf{E}\times\mathbf{B})}{(E^2+B^2)} \tag{1}$$

where $\mathbf{G} = \mathbf{S}/c^2 = (\mathbf{E}\times\mathbf{B})\,/c$ is the linear momentum density of an electromagnetic field with electric and magnetic field vectors $\mathbf{E}$ and $\mathbf{B}$ and Poynting energy flux vector $\mathbf{S}$. The $\mathcal{E} = \left(E^2+B^2\right)/2$ is the electromagnetic energy density, and c is the vacuum velocity of light.

If $\mathcal{E}/c$ and $\mathbf{G}$ were the time and space parts of a four-vector, then a Lorentz boost from the laboratory system using boost velocity $\mathbf{V} = \mathbf{u}_\mathrm{e}$ would produce a comoving reference system (denoted by primes) in which the space part of that four-vector, that is $\mathbf{G}'$ and hence the Poynting vector $\mathbf{S}' = c^2\mathbf{G}'$, would vanish, indicating a system comoving with the energy flow.[‡] Thus $\mathbf{u}_\mathrm{e}$ would be the comoving velocity of the energy flow.

However, $\mathcal{E}/c$ and $\mathbf{G}$ are *not* components of a four-vector. There is no four-vector momentum density of the form $\boldsymbol{\mathsf{G}} = (\mathcal{E}/c)\,\boldsymbol{\mathsf{e}}_0 + \mathbf{G}$.

Rather, $\mathcal{E}$ and $c\,\mathbf{G}$ are the T^{00} and T^{0i} components of the second-rank electromagnetic energy-momentum tensor

$$T^{\alpha\beta} = \begin{pmatrix} \mathcal{E} & c\,G_1 & c\,G_2 & c\,G_3 \\ c\,G_1 & M_{11} & M_{12} & M_{13} \\ c\,G_2 & M_{21} & M_{22} & M_{23} \\ c\,G_3 & M_{31} & M_{32} & M_{33} \end{pmatrix} \tag{2}$$

where $M_{ij} = -\left(E_iE_j + B_iB_j\right) + \dfrac{1}{2}\,\delta_{ij}\left(E^2+B^2\right)$.

A related point is made by Rohrlich [11], using the so-called von Laue theorem to argue that *integrals* of $\mathcal{E}/c$ and $\mathbf{G}$ over hyperplanes may in some cases transform as four-vectors. But we are treating these quantities locally, at a particular event. Von Laue's theorem does not imply that the local field functions $\mathcal{E}/c$ and $\mathbf{G}$ themselves transform as components of a four-vector. They do not. Rather than attempting to derive a four-vector from $\mathcal{E}/c$ and $\mathbf{G}$, we show how to use them in a relativistically correct manner as they are. See also Section 6.3 of [10].

Since $\mathcal{E}$ and $c\,\mathbf{G}$ are components of $T^{\alpha\beta}$, contributions to the boost transformation from the other components of $T^{\alpha\beta}$ would produce a comoving system in which $\mathbf{G}'$ and the Poynting vector would not vanish. *The electromagnetic energy flow velocity is not* $\mathbf{u}_e$.

The failure of $\mathbf{u}_\mathrm{e}$ to be the correct flow velocity can be contrasted with the well-understood theory of the flow of electric charge. The charge density ρ and the current density vector $\mathbf{J}$ are shown by the divergence of the Maxwell field tensor to form a four-vector of the form $\boldsymbol{\mathsf{J}} = c\rho\,\boldsymbol{\mathsf{e}}_0 + \mathbf{J}$. In general, $\boldsymbol{\mathsf{J}}$ can be timelike, spacelike, or null. If spacelike, there is no velocity $\mathbf{v}_\mathrm{q}$ less than the speed of light with $\mathbf{J} = \rho\,\mathbf{v}_\mathrm{q}$. But if we consider, for example, a system in which all the moving charges have the same sign, it can be shown that $\boldsymbol{\mathsf{J}}$ is timelike and hence the definition $\mathbf{u}_\mathrm{q} = \mathbf{J}/\rho$ does produce a vector of magnitude less than the speed of light. Then a Lorentz boost with boost velocity $\mathbf{V} = \mathbf{u}_\mathrm{q}$ indeed leads to a comoving primed reference system in which the current flux density $\mathbf{J}'$ vanishes[§], and $\mathbf{u}_\mathrm{q}$ is therefore the correct flow velocity of the moving charge.

But the fact that $\boldsymbol{\mathsf{J}}$ transforms as a four-vector is crucial to this argument. If it were not a four-vector transforming as in Appendix A, the system reached by boost $\mathbf{V} = \mathbf{u}_\mathrm{q}$ would have a residual current flow $\mathbf{J}' \neq 0$, and $\mathbf{u}_\mathrm{q}$ would therefore not be the correct flow velocity. The equation $\mathbf{J} = \rho\,\mathbf{u}_\mathrm{q}$ would still follow from the definition of $\mathbf{u}_\mathrm{q}$, but that formula would not imply that $\mathbf{u}_\mathrm{q}$ is the correct velocity of the flowing charge.

The failure of $\mathbf{u}_\mathrm{e}$ as a candidate for the flow velocity of electromagnetic energy is precisely because, unlike $\boldsymbol{\mathsf{J}}$, the expression written here in four-vector form $\boldsymbol{\mathsf{G}} = (\mathcal{E}/c)\,\boldsymbol{\mathsf{e}}_0 + \mathbf{G}$ actually does *not* transform as a four-vector. The equation $\mathbf{G} = \left(\mathcal{E}/c^2\right)\mathbf{u}_\mathrm{e}$ (or equivalently $\mathbf{S} = \mathcal{E}\,\mathbf{u}_\mathrm{e}$) still follows from the definition of $\mathbf{u}_\mathrm{e}$, but that formula does not imply that $\mathbf{u}_\mathrm{e}$ is the correct velocity of the flowing energy.

However, the correct boost velocity $\mathbf{V}$ can be found by starting from $\mathbf{u}_\mathrm{e}$ and applying a scalar correction factor. The corrected velocity $\mathbf{V}$ will have the same *direction* as $\mathbf{u}_\mathrm{e}$ but not the same *magnitude.* To find this corrected velocity $\mathbf{V}$ it is best to turn to a direct method, using the transformation rules for the fields $\mathbf{E}$ and $\mathbf{B}$.

---

[*] See Appendix A for a demonstration that any point at rest in the primed system moves with laboratory system coordinate velocity $\mathbf{V}$.

[†] In a discussion of the Poynting theorem in material media, but with no special attention to Lorentz covariance, Born and Wolf [3] Section 14.2, Eq. (8) identify $\mathbf{u}_\mathrm{e}$ as the *velocity of energy transport* or *ray velocity.* Section B.2 of Smith [15] echoes Born and Wolf but provides no new derivation. (The first edition of Born and Wolf's text appeared in 1959.) Geppert [5] writes a non-covariant equation with the same identification. More recently, Sebens [13,14] relies on these and other sources to identify $\mathbf{u}_\mathrm{e}$ as the electromagnetic mass flow velocity. (Following Sebens, expand $(E-B)^2 \geq 0$ and use the definitions of $\mathcal{E}$ and $\mathbf{G}$ to prove that $|\mathbf{u}_\mathrm{e}| \leq c$.)

[‡] See Appendix A for a demonstration that $\mathbf{G}'$ would be zero.

[§] Substitute $c\rho = J^0$ and $\mathbf{J}$ for G^0 and $\mathbf{G}$ in Appendix A to see that $\mathbf{J}'$ vanishes.

The rules for transformation of electric and magnetic fields by a boost with velocity $\mathbf{V}$ can be written in a special relativistically correct but not manifestly covariant form as[*]

$$\begin{aligned} \mathbf{E}' &\doteq \gamma\left(\mathbf{E}+\frac{\mathbf{V}}{c}\times\mathbf{B}\right)+(1-\gamma)\,\frac{\mathbf{V}\,(\mathbf{V}\cdot\mathbf{E})}{V^2}\\ \mathbf{B}' &\doteq \gamma\left(\mathbf{B}-\frac{\mathbf{V}}{c}\times\mathbf{E}\right)+(1-\gamma)\,\frac{\mathbf{V}\,(\mathbf{V}\cdot\mathbf{B})}{V^2} \end{aligned} \tag{3}$$

where the Lorentz factor is $\gamma=\left(1-V^2/c^2\right)^{-1/2}$. The $\doteq$ symbol means that the components of the three-vector on the left side of this symbol expressed in the primed coordinate system are numerically equal to the components of the three-vector on the right side of this symbol expressed in the original unprimed system. If $\mathbf{a}'\doteq\mathbf{c}$ and $\mathbf{b}'\doteq\mathbf{d}$, it is easily proved that: (a) $\mathbf{a}'\times\mathbf{b}'\doteq\mathbf{c}\times\mathbf{d}$ and (b) $(\mathbf{a}'\cdot\mathbf{b}')=(\mathbf{c}\cdot\mathbf{d})$.

Define the boost velocity $\mathbf{V}$ to be an unknown but rotationally scalar quantity λ times $\mathbf{u}_\mathrm{e}$

$$\mathbf{V}=\lambda\,\mathbf{u}_\mathrm{e}\,. \tag{4}$$

Since $\mathbf{u}_\mathrm{e}$ and $\mathbf{V}$ are perpendicular to both the electric and magnetic fields, it follows that $(\mathbf{V}\cdot\mathbf{E})=(\mathbf{V}\cdot\mathbf{B})=0$. Thus, (3) reduces to [†]

$$\begin{aligned} \mathbf{E}' &\doteq \gamma\left(\mathbf{E}+\frac{\mathbf{V}}{c}\times\mathbf{B}\right)\\ \mathbf{B}' &\doteq \gamma\left(\mathbf{B}-\frac{\mathbf{V}}{c}\times\mathbf{E}\right). \end{aligned} \tag{5}$$

Insert (5) into the definition $\mathbf{S}'=c\mathbf{E}'\times\mathbf{B}'$. Using (4) and then (1) leads to[‡]

$$\mathbf{S}'=c\mathbf{E}'\times\mathbf{B}'\doteq\gamma^2c\,(\mathbf{E}\times\mathbf{B})\left((u_\mathrm{e}^2/c^2)\,\lambda^2-2\lambda+1\right). \tag{6}$$

Notice that (6) verifies the statement above that $\mathbf{u}_\mathrm{e}$ is not the comoving velocity of the energy flow. Setting $\lambda=1$ in (4) makes $\mathbf{V}=\mathbf{u}_\mathrm{e}$. But setting $\lambda=1$ in (6) makes

$$\mathbf{S}'\doteq\gamma^2c\,(\mathbf{E}\times\mathbf{B})\left(u_\mathrm{e}^2/c^2-1\right)\quad\text{when}\quad\lambda=1 \tag{7}$$

which is not zero, except in the unphysical limit $u_\mathrm{e}=V=c$.

For a second and more important use of (6), choose λ to solve the quadratic equation $\left((u_\mathrm{e}^2/c^2)\lambda^2-2\lambda+1\right)=0$. Then (6) makes $\mathbf{S}'=0$. The solution is

$$\lambda=\frac{1}{(u_\mathrm{e}/c)^2}\left\{1-\sqrt{1-(u_\mathrm{e}/c)^2}\right\}. \tag{8}$$

From (4), the correct velocity of the energy flow is therefore

$$\mathbf{V}=\lambda\,\mathbf{u}_\mathrm{e}=\frac{1}{(u_\mathrm{e}/c)^2}\left\{1-\sqrt{1-(u_\mathrm{e}/c)^2}\right\}\mathbf{u}_\mathrm{e} \tag{9}$$

where $\mathbf{u}_\mathrm{e}$ is defined in (1).

This $\mathbf{V}$ is the relativistically correct boost velocity from the laboratory frame to the comoving reference frame in which $\mathbf{S}'=0$, and is therefore the laboratory system coordinate velocity of the electromagnetic energy flow[§].

Since both $\mathbf{V}$ and $\mathbf{u}_\mathrm{e}$ are parallel to the energy flux vector $\mathbf{S}$, the energy flow velocity can also be written as $\mathbf{V}=V\,(\mathbf{S}/S)$ where the magnitude V is given by[¶]

$$(V/c)=\frac{1}{(u_\mathrm{e}/c)}\left\{1-\sqrt{1-(u_\mathrm{e}/c)^2}\right\}. \tag{10}$$

This equation can be inverted to give

$$(u_\mathrm{e}/c)=\frac{2\,(V/c)}{1+(V/c)^2} \tag{11}$$

which can be used to write the correction factor λ in (8) as a function of the corrected velocity

$$\lambda=\frac{1+(V/c)^2}{2}\,. \tag{12}$$

3 Comoving energy-momentum tensor

The derivation of a reference system comoving with the flow of energy allows the electromagnetic energy-momentum tensor to be examined in more detail. The energy-momentum tensor in (2) can be transformed into the comoving (primed) coordinate system that was produced by the Lorentz boost $\mathbf{V}$. In this system, the electromagnetic energy-momentum tensor is represented by the tensor components $T'^{\alpha\beta}$ in which the cG'_i elements are zero.

$$T'^{\alpha\beta}=\begin{pmatrix} \mathcal{E}' & 0 & 0 & 0\\ 0 & M'_{11} & M'_{12} & M'_{13}\\ 0 & M'_{21} & M'_{22} & M'_{23}\\ 0 & M'_{31} & M'_{32} & M'_{33} \end{pmatrix} \tag{13}$$

where

$$\begin{aligned} \mathcal{E}' &= \frac{1}{2}\left(E'^2+B'^2\right)=\mathcal{E}\,\frac{1-(V/c)^2}{1+(V/c)^2}\\ \text{and}\quad M'_{ij} &= -\left(E'_iE'_j+B'_iB'_j\right)+\delta_{ij}\mathcal{E}'\,. \end{aligned} \tag{14}$$

We can now make another Lorentz transformation, an orthogonal spatial rotation at fixed time, to diagonalize the real, symmetric sub-matrix M_{ij} in (13).

The required spatial rotation can be defined as the product of two proper rotations. First rotate the coordinate system to bring the $\mathbf{e}'_3$ axis into the $\mathbf{V}'\doteq\mathbf{V}$ direction. Denote

---

[*] See Section 11.10 of Jackson [7], eqn (11.149).

[†] Note that $\mathbf{V}'\doteq\mathbf{V}$ as defined in Appendix A, together with (5) and property (b) of the symbol $\doteq$ noted above, imply that $(\mathbf{V}'\cdot\mathbf{E}')=\mathbf{V}\cdot\gamma\,[\mathbf{E}+(\mathbf{V}/c)\times\mathbf{B}]=\gamma(\mathbf{V}\cdot\mathbf{E})=0$. Similarly, $(\mathbf{V}'\cdot\mathbf{B}')=0$.

[‡] See a detailed derivation of (6) in Appendix B.

[§] Appendix C gives details of the comoving system for possible values of $(\mathbf{E}\cdot\mathbf{B})$ at a given event.

[¶] Footnote † on page 4 proves that $0\leq u_\mathrm{e}\leq c$. As u_e/c increases from 0 to 1, (10) shows that V/c increases monotonically from 0 to 1, with $V\leq u_\mathrm{e}$ at every point. It follows that $0\leq V\leq c$ also.

this rotated system by tildes. Rotations do not change three-vectors, which are invariant objects under rotations. However, rotations do change the *components* of three-vectors. Thus $\tilde{\mathbf{V}} = \mathbf{V}'$, $\tilde{\mathbf{E}} = \mathbf{E}'$, and $\tilde{\mathbf{B}} = \mathbf{B}'$, but in the tilde system $\tilde{\mathbf{V}}$ now has components $\tilde{V}_1 = \tilde{V}_2 = 0$ and $\tilde{V}_3 = V$. Then using Footnote † on page 5, we have $0 = (\mathbf{E}' \cdot \mathbf{V}') = (\tilde{\mathbf{E}} \cdot \tilde{\mathbf{V}}) = V\tilde{E}_3$. Since the magnitude $V \neq 0$, it follows that $\tilde{E}_3 = 0$. A similar argument proves that $\tilde{B}_3 = 0$. Thus the {33} component of the energy-momentum tensor when expressed in the tilde system is $\tilde{T}^{33} = -\left(\tilde{E}_3^2 + \tilde{B}_3^2\right) + \tilde{\mathcal{E}} = \tilde{\mathcal{E}}$. The tensor from (13), when expressed in the tilde system, becomes

$$\tilde{T}^{\alpha\beta} = \begin{pmatrix} \tilde{\mathcal{E}} & 0 & 0 & 0 \\ 0 & \tilde{M}_{11} & \tilde{M}_{12} & 0 \\ 0 & \tilde{M}_{21} & \tilde{M}_{22} & 0 \\ 0 & 0 & 0 & \tilde{\mathcal{E}} \end{pmatrix} \tag{15}$$

where $\tilde{\mathcal{E}} = \mathcal{E}'$.

Since the invariant trace of the electrodynamic energy-momentum tensor vanishes[*], it follows from (15) that

$$0 = \eta_{\alpha\beta}\tilde{T}^{\alpha\beta} = -\tilde{\mathcal{E}} + \tilde{M}_{11} + \tilde{M}_{22} + \tilde{\mathcal{E}} \tag{16}$$

and hence $\tilde{M}_{11} = -\tilde{M}_{22}$. Also, the symmetry of the energy-momentum tensor makes $\tilde{M}_{21} = \tilde{M}_{12}$. Thus

$$\tilde{T}^{\alpha\beta} = \begin{pmatrix} \tilde{\mathcal{E}} & 0 & 0 & 0 \\ 0 & -\tilde{\psi} & \tilde{\xi} & 0 \\ 0 & \tilde{\xi} & \tilde{\psi} & 0 \\ 0 & 0 & 0 & \tilde{\mathcal{E}} \end{pmatrix} \tag{17}$$

where $\tilde{\psi} = \tilde{M}_{22}$ and $\tilde{\xi} = \tilde{M}_{12}$.

A second proper rotation, this time about the $\tilde{\mathbf{e}}_3$ axis, produces the final coordinate system, denoted with double primes. After this rotation, $E''_3 = \tilde{E}_3 = 0$, $B''_3 = \tilde{B}_3 = 0$, and $\mathbf{V}'' = \tilde{\mathbf{V}}$ has components $V''_1 = V''_2 = 0$ and $V''_3 = V$. The only effect of this second rotation is to diagonalize the matrix $\begin{pmatrix} -\tilde{\psi} & \tilde{\xi} \\ \tilde{\xi} & \tilde{\psi} \end{pmatrix}$. The energy momentum tensor then has its final, diagonal form in the double-prime system

$$T''^{\alpha\beta} = \begin{pmatrix} \mathcal{E}'' & 0 & 0 & 0 \\ 0 & -a'' & 0 & 0 \\ 0 & 0 & a'' & 0 \\ 0 & 0 & 0 & \mathcal{E}'' \end{pmatrix} \tag{18}$$

where $\mathcal{E}'' = \tilde{\mathcal{E}} = \mathcal{E}'$. The parameter a'' has absolute value $|a''| = \left\{\tilde{\psi}^2 + \tilde{\xi}^2\right\}^{1/2}$ where $\pm\left\{\tilde{\psi}^2 + \tilde{\xi}^2\right\}^{1/2}$ are the two eigenvalues of the matrix $\begin{pmatrix} -\tilde{\psi} & \tilde{\xi} \\ \tilde{\xi} & \tilde{\psi} \end{pmatrix}$ that were calculated during the diagonalization process. The sign of a'' depends on the directions and relative magnitudes of the electric and magnetic fields.

The rotation that takes the system from the primed to the double-primed system is then the product of the first and second rotations. The various representations of the boost velocity are related by $\mathbf{V}'' = V\mathbf{e}''_3 = \tilde{\mathbf{V}} = V\tilde{\mathbf{e}}_3 = \mathbf{V}' \overset{\circ}{=} \mathbf{V}$ where all of these vectors have the same original magnitude V.

The energy-momentum tensor in the double-prime system is diagonal and in a canonical form that depends only on the energy density $\mathcal{E}''$ in the comoving system and one other parameter a''.

Section 4 shows that there are realistic electromagnetic cases in which $a'' \neq 0$ and hence the diagonal elements M''_{ii} for $i = 1, 2, 3$ are not all equal, unlike the analogous elements in the energy-momentum tensor of a perfect fluid[†], all of which are equal by definition, a fact of relevance for future studies that might attempt a fluid-dynamic model of electrodynamic energy flow.

4 Example: crossed static fields

Consider an example with static, perpendicular electric and magnetic fields.[‡] Choose the Cartesian axes of the laboratory system so that $\mathbf{E} = E\,\hat{\mathbf{x}}$ and $\mathbf{B} = B\,\hat{\mathbf{y}}$. Then (1) becomes

$$\frac{\mathbf{u}_\mathrm{e}}{c} = \left(\frac{2E\,B}{E^2 + B^2}\right)\hat{\mathbf{z}}\,. \tag{19}$$

The energy flow velocity is thus $\mathbf{V} = V\,\hat{\mathbf{z}}$ where $V = \lambda\,u_\mathrm{e}$ with λ from (8). Inserting this $\mathbf{V}$ into (5) with the above values of the electric and magnetic fields gives

$$\begin{aligned} \mathbf{E}'' &\overset{\circ}{=} \gamma\left(E - \frac{VB}{c}\right)\hat{\mathbf{x}} \\ \mathbf{B}'' &\overset{\circ}{=} \gamma\left(B - \frac{VE}{c}\right)\hat{\mathbf{y}}\,. \end{aligned} \tag{20}$$

Thus the definitions in (2) when applied in the double-prime system give $M''_{ij} = 0$ for $i \neq j$ and

$$\begin{aligned} M''_{11} &= -E''^2_1 + \mathcal{E}'' = -\frac{1}{2}\left(E''^2_1 - B''^2_2\right) \\ M''_{22} &= -B''^2_2 + \mathcal{E}'' = \frac{1}{2}\left(E''^2_1 - B''^2_2\right) \\ M''_{33} &= \mathcal{E}'' = \frac{1}{2}\left(E''^2_1 + B''^2_2\right)\,. \end{aligned} \tag{21}$$

where E''_1 and B''_2 are the components of $\mathbf{E}''$ and $\mathbf{B}''$, respectively, in (20).

The step of rotating from primed to double-primed reference systems that was necessary to move from (13) to (18) above was not necessary here due to a propitious choice of original laboratory reference system. The Lorentz boost with velocity $\mathbf{V} = V\,\hat{\mathbf{z}}$ produces an already diagonal energy momentum tensor with $M''_{ij} = 0$ for $i \neq j$.

---

[*] See Section 7.8 of Rindler [9].

[†] See Part I, Chapter 2, Section 10 of Weinberg [16].

[‡] The center of a parallel plate capacitor at the center of a long solenoid, for example.

Consider the case $E \neq B$. From (19), this inequality implies that $u_e < c$ and hence, from (10), that $V < c$, a physically possible value. Also $E \neq B$ implies, either from the invariance of $\left(E^2 - B^2\right)$ noted in Appendix C or directly from (20), that $2M''_{22} = \left(E''^2_1 - B''^2_2\right) = \left(E^2 - B^2\right) \neq 0$. Thus $E \neq B$ implies that $M''_{11} = -M''_{22} \neq 0$ and hence that $M''_{11} \neq M''_{22}$.

Comparing (21) to (14) and (18) shows that for the crossed-field example with $E{\neq}B$, the energy-momentum tensor in the double-prime system is (18) with

$$\begin{aligned} a'' &= \left(E^2 - B^2\right)/2 \neq 0 \\ \text{and} \quad \mathcal{E}'' &= \mathcal{E}\left(1 - (V/c)^2\right)/\left(1 + (V/c)^2\right) \end{aligned} \tag{22}$$

where $\mathcal{E} = \left(E^2 + B^2\right)/2$.

As asserted at the end of Section 3, the inequality $E \neq B$ in the crossed-field example shows a physically reasonable case for which $a'' \neq 0$ and the M''_{ii} for $i = 1, 2, 3$ are not all equal.[*]

There are questions about the interpretation of this example globally[†]. In our use of this example, however, we need not consider the question of so-called *hidden momentum* required to balance the total field momentum[‡]. Here, the only relevant use of this example is to illustrate the correct local definition of the energy flow velocity and comoving energy-momentum tensor in a vacuum region where the fields are well known — at the center of the parallel plate capacitor far from the edges.

5 Relativistic mass density

The energy density $\mathcal{E}$ of either static or time-varying vacuum electromagnetic fields can be used to define a *relativistic* mass density[§]

$$\mathcal{M}_{\text{rel}} = \mathcal{E}/c^2\,. \tag{23}$$

The adjective *relativistic* indicates that this mass density is analogous to a single-particle relativistic mass $m_{\text{rel}} = \gamma m = e/c^2$ where e is the particle relativistic energy, m is the invariant or rest mass of the particle, and $\gamma = \left(1 - v^2/c^2\right)^{-1/2}$ is the Lorentz factor of its velocity $\mathbf{v}$.

It follows from (23) that the flow velocity of the energy $\mathcal{E}$, the velocity $\mathbf{V}$ derived in Section 2 and summarized in (9), must also be the flow velocity of the relativistic mass $\mathcal{M}_{\text{rel}}$.

In the single-particle case, the same m_{rel} can be used to relate the momentum of the particle to its velocity, $\mathbf{p} = \gamma m \mathbf{v} = m_{\text{rel}}\mathbf{v}$. However in the case of electromagnetic fields, the same mass density $\mathcal{M}_{\text{rel}}$ cannot be used for both purposes.

Due to the correction of the flow velocity in Section 2, which was necessitated by adherence to the transformation rules of special relativity, the relation between momentum density $\mathbf{G}$ and corrected flow velocity $\mathbf{V}$ is

$$\mathbf{G} = \left(\frac{\mathcal{M}_{\text{rel}}\,\mathbf{V}}{\lambda}\right) \neq \mathcal{M}_{\text{rel}}\mathbf{V} \tag{24}$$

where λ is the correction factor in (12).

The inequality in (24) shows that the electromagnetic momentum density at an event is not equal to the electromagnetic mass density at that event times the relativistically correct mass flow velocity there.

The explicit expression for the correction factor λ from (12) quantifies the extent of the inequality. The effective mass for momentum calculation is the larger value $\mathcal{M}_{\text{rel}}/\lambda$ rather than $\mathcal{M}_{\text{rel}}$[¶].

The failure of $\mathcal{M}_{\text{rel}}\mathbf{V}$ to equal the momentum density $\mathbf{G}$ in (24) suggests that vacuum electromagnetic field momentum cannot be explained only by the motion of relativistic mass. There must be another source of real electromagnetic field momentum.

6 Conclusion

The electromagnetic field contains energy and momentum. Calculation of the energy flow velocity and energy-momentum tensor in a relativistically correct manner opens the subject to new insights into that energy and momentum. For example, the energy-momentum tensor measured by an observer comoving with the flow velocity is obtained in diagonal, canonical form suggestive of possible fluid dynamical models. And the momentum density of the electromagnetic field is shown to require some source other than the flow of relativistic mass.

Appendix A: Lorentz boosts

Consider a Lorentz transformation from an *unprimed* coordinate system (which we also refer to as the *laboratory* system) with coordinates $x = (x^0, x^1, x^2, x^3)$ to a *primed* coordinate system with coordinates $x' = (x'^0, x'^1, x'^2, x'^3)$ where $x^0 = ct$ and $x'^0 = ct'$. The most general proper, homogeneous Lorentz transformation from the unprimed to the primed systems can be written as a Lorentz boost times a rotation.[‖]

Definition of Lorentz boost

A Lorentz boost transformation is parameterized by a boost velocity vector $\mathbf{V}$ with components (V_1, V_2, V_3) and magnitude $V = \left(V_1^2 + V_2^2 + V_3^2\right)^{1/2}$. Using the Einstein summation

---

[*] Our use of this example is based on $E \neq B$. The case $E = B \neq 0$ would have to be approached as a limit, as discussed in Appendix C (c). With $E \neq 0$ and $B = E(1 + \delta)$, retaining leading order in the small quantity δ gives $(u_e/c) \approx (1 - \delta^2/2)$, $\lambda \approx (1 - |\delta|)$, $(V/c) \approx (1 - |\delta|)$, $(\mathcal{E}''/E^2) \approx |\delta|$, $(T''^{\alpha\beta}/E^2) \approx \operatorname{diag}(|\delta|, \delta, -\delta, |\delta|)$, and $\left(a''/E^2\right) \approx -\delta$.

[†] See McDonald [8] for calculation of the total field momentum of a similar example.

[‡] See, for example, Babson et al [1].

[§] For example, see Section 3 of Sebens [13].

[¶] Note that (24) can be written as $\mathbf{G} = \mathcal{M}_{\text{eff}}\mathbf{V}$ where, using (12), (14), and (23), $\mathcal{M}_{\text{eff}} = \mathcal{M}_{\text{rel}}/\lambda = 2\gamma^2(\mathcal{E}'/c^2)$ where $\gamma = (1 - V^2/c^2)^{-1/2}$.

[‖] See Part I, Chapter 2, Section 1 of Weinberg [16].

convention, it is written as $x'^{\alpha} = \Lambda^{\alpha}_{\ \beta}\, x^{\beta}$ where $\Lambda^{0}_{\ 0} = \gamma$, $\Lambda^{0}_{\ i} = \Lambda^{i}_{\ 0} = -\gamma V_i/c$, and $\Lambda^{i}_{\ j} = \delta_{ij} + (\gamma - 1)V_iV_j/V^2$. The δ_{ij} is the Kronecker delta function. Also $\gamma = \left(1 - V^2/c^2\right)^{-1/2}$.

The inverse boost $\underline{\Lambda}^{\alpha}_{\ \beta}$ is the same except for the substitution $V_i \to -V_i$. Thus the inverse boost vector is $(-\mathbf{V}')$ where $\mathbf{V}' \overset{\circ}{=} \mathbf{V}$.

Meaning of the boost velocity V

The velocity $\mathbf{V}$ that parameterizes the Lorentz boost is also the coordinate velocity, as measured from the unprimed laboratory system, of any point that is at rest in the primed system. In this sense, the entire primed system is moving with velocity $\mathbf{V}$ as observed from the laboratory system.

To see this, apply the inverse Lorentz boost to the differentials of a point at rest in the primed system, $dx'^{i} = 0$ for $i = 1, 2, 3$, but $dx'^{0} > 0$. The result is $dx^0 = \gamma dx'^0$ and $dx^i = \gamma\,(V_i/c)\,dx'^0$. Thus $dx^i/dt = V_i$, as was asserted.

Consequence of existence of a four-vector G

The discussion surrounding (2) shows that $\mathsf{G} = (\mathcal{E}/c)\,\mathsf{e}_0 + \mathbf{G}$ is *not* a four-vector, despite being written in four-vector form here. Its components instead transform as components of the energy-momentum tensor. But suppose for a moment that it is a four-vector. If so, then a Lorentz boost with a boost velocity $(\mathbf{u}_{\rm e}/c) = \mathbf{G}/\,(\mathcal{E}/c)$ would make the transformed space part of G equal to zero.

As applied to a four-vector, the Lorentz boost transformation rule is $G'^{\alpha} = \Lambda^{\alpha}_{\ \beta}\, G^{\beta}$. Hence

$$\begin{aligned} G'^{i} &= \Lambda^{i}_{\ 0}G^0 + \Lambda^{i}_{\ j}G^j \\ &= -\gamma\frac{V_i}{c}G^0 + G^i + (\gamma - 1)\frac{V_i\left(V_jG^j\right)}{V^2}\,. \end{aligned} \tag{25}$$

Replacing V_i/c by $(u_{\rm e})_i\,/c = G^i/G^0$ in (25) makes $G'^{i} = 0$, as asserted.

Appendix B: Detailed derivation of Eq. (6) for S′

We have (1), (4) and (5) and $(\mathbf{V}\cdot\mathbf{E}) = (\mathbf{V}\cdot\mathbf{B}) = 0$. Inserting (5) into $\mathbf{S}' = c\,(\mathbf{E}'\times\mathbf{B}')$ gives

$$\mathbf{S}' = c\mathbf{E}'\times\mathbf{B}' \overset{\circ}{=} c\gamma^2\left\{(\mathbf{E}\times\mathbf{B}) + \mathbf{f} + \mathbf{g}\right\} \tag{26}$$

where, omitting zero terms,

$$\begin{aligned} \mathbf{f} &= -\mathbf{E}\times\left(\frac{\mathbf{V}}{c}\times\mathbf{E}\right) + \left(\frac{\mathbf{V}}{c}\times\mathbf{B}\right)\times\mathbf{B} \\ &= -\left(E^2 + B^2\right)\frac{\mathbf{V}}{c} = -\lambda\left(E^2 + B^2\right)\frac{\mathbf{u}_{\rm e}}{c} \\ &= -\lambda\left(E^2 + B^2\right)\frac{2\,(\mathbf{E}\times\mathbf{B})}{(E^2 + B^2)} = -2\lambda\,(\mathbf{E}\times\mathbf{B}) \end{aligned} \tag{27}$$

and, again omitting zero terms,

$$\begin{aligned} \mathbf{g} &= -\left(\frac{\mathbf{V}}{c}\times\mathbf{B}\right)\times\left(\frac{\mathbf{V}}{c}\times\mathbf{E}\right) \\ &= -\frac{\mathbf{V}}{c}\left\{\left(\frac{\mathbf{V}}{c}\times\mathbf{B}\right)\cdot\mathbf{E}\right\} = \frac{\mathbf{V}}{c}\left\{\frac{\mathbf{V}}{c}\cdot(\mathbf{E}\times\mathbf{B})\right\} \\ &= \lambda^2\left\{\frac{2\,(\mathbf{E}\times\mathbf{B})}{(E^2 + B^2)}\right\}\left\{\frac{\mathbf{u}_{\rm e}}{c}\cdot\left(\frac{E^2 + B^2}{2}\right)\frac{\mathbf{u}_{\rm e}}{c}\right\} \\ &= \lambda^2\left(\frac{\mathbf{u}_{\rm e}}{c}\cdot\frac{\mathbf{u}_{\rm e}}{c}\right)(\mathbf{E}\times\mathbf{B}) = \lambda^2\left(\frac{u_{\rm e}}{c}\right)^2(\mathbf{E}\times\mathbf{B})\,. \end{aligned} \tag{28}$$

Collect terms and factor out $\mathbf{E}\times\mathbf{B}$ to get

$$\mathbf{S}' = c\mathbf{E}'\times\mathbf{B}' \overset{\circ}{=} \gamma^2 c\,(\mathbf{E}\times\mathbf{B})\left\{\left(\frac{u_{\rm e}}{c}\right)^2\lambda^2 - 2\lambda + 1\right\} \tag{29}$$

which is (6).

Appendix C: Detail of the comoving system

The comoving system is defined by $\mathbf{S}' = c\,(\mathbf{E}'\times\mathbf{B}') = 0$. Thus $|\mathbf{E}'\times\mathbf{B}'| = E'B'\sin\theta' = 0$ where θ' is the angle between $\mathbf{E}'$ and $\mathbf{B}'$ in the comoving system.

From Eqs (7.62) and (7.63) of Rindler [9], we have $\left(E'^2 - B'^2\right) = \left(E^2 - B^2\right)$ and $(\mathbf{E}'\cdot\mathbf{B}') = (\mathbf{E}\cdot\mathbf{B})$. It follows that:

(a) An event with $\mathbf{E}\cdot\mathbf{B} \neq 0$ has $E'B' \neq 0$ and therefore $\mathbf{E}'$ and $\mathbf{B}'$ must be either parallel or anti-parallel, $\theta' = 0$ or $\theta' = \pi$, at this event;

(b) An event with $0 = (\mathbf{E}\cdot\mathbf{B}) = (\mathbf{E}'\cdot\mathbf{B}') = E'B'\cos\theta'$ cannot have $E'B' \neq 0$ in the comoving system because that would require both $\cos\theta' = 0$ and $\sin\theta' = 0$. Thus $E'B' = 0$ and one of E' and B' must be zero. If $E > B$ then $E' > B'$ and hence $B' = 0$. If $E < B$ then $E' < B'$ and hence $E' = 0$;

(c) If both $0 = \mathbf{E}\cdot\mathbf{B}$ and $E = B \neq 0$ at an event, then both $E'B' = 0$ and $E' = B'$, and therefore $E' = B' = 0$. But (1) and (10) show that such an event also has $u_{\rm e}/c = 1$ and hence $V/c = 1$. The case $E = B \neq 0$ and $0 = \mathbf{E}\cdot\mathbf{B}$ therefore must be approached as a limit.

Received on Sept. 23, 2020

References

1. Babson D., Reynolds S. P., Bjorkquist R., Griffiths D. J. Hidden momentum, field momentum, and electromagnetic impulse. *American Journal of Physics*, 2009, v. 77, 826–833.
2. Boos F. L. More on the Feynman's disk paradox. *American Journal of Physics*, 1984, v. 52 (8), 756.
3. Born M., Wolf E. Principles of Optics, 6th edn. Pergamon Press, Oxford, 1980.
4. Feynman R., Leighton R., Sands M. The Feynman Lectures on Physics, vol II. Addison-Wesley Pub. Co., Reading, MA, 1964.
5. Geppert D. V. Energy-transport velocity in electromagnetic waves. *Proceedings of the IEEE (correspondence)*, 1965, v. 53, 1790.
6. Griffiths D. J. Introduction to Electrodynamics, 3rd edn. Prentice Hall, Upper Saddle River, NJ, 1999.

7. Jackson J. D. Classical Electrodynamics, 2nd edn. John Wiley and Sons, New York, 1975.
8. McDonald K. T. Electromagnetic momentum of a capacitor in a uniform magnetic field. 2015. https://www.hep.princeton.edu/~mcdonald/examples/cap_momentum.pdf
9. Rindler W. Relativity Special, General, and Cosmological, 2nd edn. Oxford University Press, Oxford, 2006.
10. Rohrlich F. Classical Charged Particles. Addison-Wesley Pub. Co., Reading, MA, 1965.
11. Rohrlich F. Answer to question 26 "Electromagnetic field momentum" *American Journal of Physics*, 1996, v, 64, 16.
12. Romer R. H. Angular momentum of static electromagnetic fields. *American Journal of Physics*, 1966, v. 34, 772.
13. Sebens C. T. Forces on fields. *Studies in the History and Philosophy of Modern Physics*, 2018, v. 63, 1–11.
14. Sebens C. T. The mass of the gravitational field. *British Journal for the Philosophy of Science*, forthcoming. arXiv: abs/1811.10602v2.
15. Smith R. L. The velocities of light. *American Journal of Physics*, 1970, v. 38,(8), 978–984.
16. Weinberg S. Gravitation and Cosmology. John Wiley and Sons, New York, 1972.

Plausible Fundamental Origins of Emissivity (I)

S. Msindo[†*], G. G. Nyambuya[‡], and C. Nyamhere[†]

[‡]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, Zimbabwe.
[†]Midlands State University – Faculty of Science and Technology,
Department of Applied Physics and Telecommunications, P. Bag 9055, Gweru, Zimbabwe.
[*]E-mail: simbarashemsindo@gmail.com

Emissivity is a fundamental property of matter that measures the ratio of the thermal radiation emanating from a thermodynamic surface to the radiation from an ideal black-body surface at the same temperature and it takes values from 0 to 1. This property is not a theoretically derived thermodynamic property of matter, but a *posteriori justified* property that is derived from experiments after its need was found necessary in order to balance up the theoretically expected radiation of a black-body at the same temperature to that actually measured in the laboratory for a material body at the same temperature. From a fundamental theoretical stand-point, we argue herein that emissivity may arise due perhaps to the *existence of non-zero finite lower and upper cut-off frequencies* in the thermal radiation of matter, thus leading to material bodies emitting not all the radiation expected from them when compared to equivalent black-body surfaces. We demonstrate that a non-zero lower limiting frequency is implied by the refractive index of materials, while an upper limit frequency is adopted from Debye's (1912) ingenious idea of an upper limiting cut-off frequency which arises from the fact that the number of modes of vibrations of a finite number of oscillators must be finite.

1 Introduction

Emissivity is a fundamental, intrinsic and inherent property of all known materials. Commonly, one talks of the emissivity of solid materials and as such, emissivity is a property typically associated with solids. In reality, all forms (solid, liquid, gas) of matter exhibit this property. In general, the emissivity of a given material is defined as the ratio of the thermal radiation from a surface to the radiation of an ideal black surface at the same temperature. As presently obtaining, this important property of matter – *emissivity* – has no fundamental theoretical justification – it is an experimentally derived property of matter. This article seeks to lay down a theoretical framework and basis that not only justifies the existence of this property of matter, but to investigate this from a purely theoretical standpoint.

To that end, in the present article, we conduct an *initial forensic analysis* of the modern derivation of the Planck Radiation Law (PRL) [1–3]. In this analysis, we identify two *loopholes* in the derivation of the PRL, and these are:

1. **Dispersion Relation Problem:** The dispersion relation assumed in the PRL is that of a photon in a *vacuo*, i.e.:

$$E = pc_0, \tag{1}$$

where: (E, p) are the (kinetic) energy and momentum of the photon in question, and: $c_0 = 2.99792458 \times 10^8\,\mathrm{m\,s^{-1}}$ is the speed of light in *vacuo* (2018 CODATA[*]). This Eq. (1), is what is used in the derivation of the PRL in relation to the energy and momentum of the photon in the interior of material bodies. Without an iota of doubt, the interior of material bodies is certainly not a *vacuo*. This means that the dispersion relation (1) is not the appropriate dispersion relation to describe these photons generated therein material bodies. We need to use the correct equation – i.e. by replacing (1) with:

$$E = pc, \tag{2}$$

where: $c = c_0/n$; and here: c, is the speed of light in the material (medium) whose refractive index is n and n: $0 < n < 1$. This is the first correction to the PRL that we shall conduct.

2. **Limits Problem:** The second correction has to do with the lower and upper limits in the integral leading to the PRL. As one will notice (and most probably ignore) is that the derivation leading to the PRL does not have a finite upper limit (i.e. $\nu_{\mathrm{H}} = \infty$) and at the same time, this same integral has a lower bound limit of zero (i.e. $\nu_{\mathrm{L}} = 0$). What this means is that the photons emitted by material bodies have wavelengths in the range – zero ($\nu_{\mathrm{L}} = 0$) to infinity ($\nu_{\mathrm{H}} = \infty$). A zero frequency photon implies zero kinetic energy and an infinite frequency photon implies an infinite kinetic energy of the photon. The lower bound frequency ($\nu_{\mathrm{L}} = 0$) has serious problems with Heisenberg's uncertainty principle [4], while the upper infinite frequency ($\nu_{\mathrm{H}} = \infty$) has obvious *topological defects* with physical and natural reality as we know it.

Using the above *two points of critique* in the derivation of the PRL, we shall advance a thesis which seeks to demonstrate that, it is possible *in principle* to justify from a physical and fundamental theoretical level the existence and the need of the emissivity function of a material. There is no such effort in the present literature where such an endeavour has been attempted – this, at least is our view point derived from the wider literature that we have managed to lay our hands on.

---

[*] https://physics.nist.gov/cgi-bin/cuu/Value?c

Now, in closing this introductory section, we shall give a synopsis of the remainder of this article. In §2 and §3, for self-containment, instructive and completeness purposes, we present an exposition of the Planck radiation theory and the derivation of the Stefan-Boltzmann Law respectively, where emphasis is made on the two points of critique to the Planck theory that we made above. In §4, we present our derivation. In §5, a general discussion is presented. Lastly, in §6, in a rather succinct manner, the conclusion drawn from the present work is laid down.

2 Planck radiation theory

As was presented in the first article of this series [5], we shall make the derivation of the PRL our point of departure. We know that the number of quantum states dN in the momentum volume space d^3p and physical volume space V, is given by:

$$dN = \frac{2Vd^3p}{h^3}, \tag{3}$$

where: $h = 6.62607015 \times 10^{-34}$ J s is Planck's constant (2018 CODATA[*]). The factor 2 in (3) comes in because of the number of degrees of freedom of the photon: one for traverse and the other for longitudinal polarisation – i.e. the photon has two polarization states. Now, given that: $d^3p = 4p^2dp$, it follows that:

$$dN = \frac{8\pi V p^2 dp}{h^3}, \tag{4}$$

and further, given that for a photon of momentum p_γ, energy E_γ and frequency ν, its energy-momentum is such that: $p_\gamma = E_\gamma/c_0 = h\nu/c_0$, it follows from this, that the number of modes in the frequency interval: ν to $\nu + d\nu$ is:

$$dN = \left(\frac{8\pi V}{c_0^3}\right)\nu^2 d\nu. \tag{5}$$

The actual number of occupied states dn is such that $dn = f_{\mathrm{BE}}(\nu, \mathcal{T})\, dN$ where:

$$f_{\mathrm{BE}}(\nu, \mathcal{T}) = \frac{1}{e^{h\nu/k_{\mathrm{B}}T} - 1}, \tag{6}$$

is the Bose-Einstein probability function which for a temperature $\mathcal{T}$, gives the probability of occupation of a quantum state whose energy is $E_\gamma = h\nu$ and: $k_{\mathrm{B}} = 1.38064852(79) \times 10^{-23}$ J K^{-1} is the Boltzmann constant (2018 CODATA[†]). From the foregoing:

$$dn_* = \frac{8\pi V}{c_0^3}\,\frac{\nu^2 d\nu}{e^{h\nu/c_0 h_{\mathrm{B}}T} - 1}, \tag{7}$$

leading to the energy density: $B_\nu(\nu, T)d\nu = E_\gamma dn_*/V$, now being given by:

$$B_\nu(\nu, T)d\nu = \frac{8\pi h}{c_0^3}\,\frac{\nu^3 d\nu}{e^{h\nu/k_{\mathrm{B}}T} - 1}, \tag{8}$$

where: $B_\nu(\nu, T)$ is the spectral irradiance given in terms of ν: (8) is our sought-for PRL.

3 Stephan-Boltzmann law

Now, to derive the Stefan-Boltzmann Law (SBL) from (8), we start by setting: $x = h\nu/k_{\mathrm{B}}T$. This setting implies that: $d\nu = k_{\mathrm{B}}T dx/h$, thus substituting this into (8), we then have:

$$B_\nu(\nu, T)d\nu = \frac{8\pi k_{\mathrm{B}}^4 \mathcal{T}^4}{h^3 c_0^3}\,\frac{x^3 dx}{e^x - 1}. \tag{9}$$

From the foregoing theory, the total energy density $\mathscr{E}_{\mathrm{theo}}$ radiated per unit time by a radiating body is such that:

$$\begin{aligned} \mathscr{E}_{\mathrm{theo}} &= \frac{c_0}{4}\int_{\nu_{\mathrm{L}}=0}^{\nu_{\mathrm{H}}=\infty} B_\nu(\nu, T)d\nu, \\ &= \frac{2\pi k_{\mathrm{B}}^4 \mathcal{T}^4}{h^3 c_0^2}\int_0^\infty \frac{x^3 dx}{e^x - 1}, \end{aligned} \tag{10}$$

and given that: $\int_0^\infty x^2 dx/(e^x - 1) = \pi^4/15$, it follows that the SBL will thus be given by:

$$\mathscr{E}_{\mathrm{theo}} = \sigma_0 \mathcal{T}^4, \tag{11}$$

where one can most easily deduce that the fundamental and universal constant – the *Stefan-Boltzmann constant*: $\sigma_0 = 2\pi^2 k_{\mathrm{B}}^4/15h^3c_0^2$. In terms of its actually experimentally measured value: $\sigma_0 = 5.670374419 \times 10^{-8}$ W m^{-2} K^{-4} (2018 CODATA[‡]).

Written as it appears in (11), the SBL is not compatible with physical and natural reality as it needs to be supplemented with a new term – namely the emissivity ϵ, i.e.:

$$\mathscr{E}_{\mathrm{exp}} = \epsilon\sigma_0\mathcal{T}^4. \tag{12}$$

The above result is what one gets from experiments. We shall derive the emissivity function: $\epsilon = \epsilon(\nu, \mathcal{T})$ from the fundamental soils of theory.

4 Derivation

In this section, we shall in two parts, i.e. §4.1 and §4.2, derive a relation that connects the emissivity function with the refractive index of the given material and both the upper and lower limits in the energy of the photon.

4.1 Dispersion relation problem

In the derivation of the PRL, i.e. (8), and as well as the SBL, i.e. (11), we have used the *vacuo* dispersion relation (1) for the photon. As stated in the introductory section, this is *not correct* as one is supposed to use the correct non-*vacuo* photon dispersion relation (2). If we do the *correct* thing and

---

[*] https://physics.nist.gov/cgi-bin/cuu/Value?h

[†] https://physics.nist.gov/cgi-bin/cuu/Value?k

[‡] https://physics.nist.gov/cgi-bin/cuu/Value?sigma

instead use (2) in the derivation of the PRL, instead of the PRL given in (8), we will obtain the new revised PRL:

$$\begin{aligned} B_\nu(\nu,T)d\nu &= \frac{8\pi h}{c^3}\frac{\nu^3 d\nu}{e^{h\nu/k_\mathrm{B}T}-1}, \\ &= \frac{8\pi h}{c_0^3}\frac{n^3\nu^3 d\nu}{e^{h\nu/k_\mathrm{B}T}-1}. \end{aligned} \tag{13}$$

The difference between (13) and (8), is the introduction of the refractive index, n.

Now, from this new PRL (13) together with the correct non-*vacuo* photon dispersion relation (2), one obtains the following refractive index modified SBL:

$$\mathscr{E}_\mathrm{exp} = \frac{c_0}{4}\int_{\nu_\mathrm{L}}^{\nu_\mathrm{H}} \frac{cB_\nu(\nu,T)d\nu}{c_0} = \frac{c_0}{4}\int_{\nu_\mathrm{L}}^{\nu_\mathrm{H}} \frac{B_\nu(\nu,T)d\nu}{n}, \tag{14}$$

where in (14), we have not set the limits ($\nu_\mathrm{H} = \infty; \nu_\mathrm{L} = 0$), but have left this as a task to be dealt with in §4.2.

Now – proceeding to institute in (14) the substitution: $x = h\nu/k_\mathrm{B}T$, and remembering that the refractive index n is a function of ν and possibly $\mathcal{T}$ as well (i.e. $n = n(\nu,\mathcal{T}) = n(x)$), it follows that (14) will reduce to:

$$\mathscr{E}_\mathrm{exp} = \sigma_0\mathcal{T}^4 \int_{x_\mathrm{L}}^{x_\mathrm{H}} \frac{15}{\pi^4}\frac{x^3 n^2(x)dx}{e^x-1}. \tag{15}$$

With $\mathscr{E}_\mathrm{exp}$ now written as it has been written in (15), one can reasonably identify the emissivity function as:

$$\epsilon = \frac{15}{\pi^4}\int_{x_\mathrm{L}}^{x_\mathrm{H}} \frac{x^3 n^2(x)dx}{e^x-1} = \epsilon(x) = \epsilon(\nu,\mathcal{T}). \tag{16}$$

In this way, the emissivity has not been introduced as a result of an experimental requirement, but foisted by subtle theoretical requirements to do with the (obvious *but* neglected) shortcomings stated in the introduction section.

Our intention in the present article is not to investigate this newly-derived emissivity function (16), but merely to make a statement to the effect that the emissivity function can be derived from the fundamental soils of theoretical physics. We shall slate for the next installation, the task to test the emissivity function (16) against real data. In the subsequent subsection, we will now deal with the issue of the limits in the integral (16).

4.2 Limits problem

As stated previously, a photon frequency of zero (i.e. photon with zero energy) does not make sense especially in the face of Heisenberg's [4] uncertainty limit. To obtain a reasonable estimate of this, one can appeal to logical reasoning by simply asking the question: *What is the largest wavelength of a photon that can travel in a medium with a mean inter-molecular spacing: $\ell = \ell(T)$?* We know that the speed of our photon is c and that this speed is such that it is equal to: $\lambda\nu$, where λ is the wavelength of our photon. In order for the smooth passage of the photon in such a medium, it is reasonable to assume that the wavelength of the photon be at most equal to one half of the mean spacing of the given medium, i.e. $\lambda_\mathrm{max} = \ell/2$. Given that: $c = \lambda\nu$, it follows that we must have: $\nu_\mathrm{L} = 2c_0/n\ell$.

Now, in establishing the upper limiting frequency that must enter the integral leading to the PRL, we will use the reasoning already laid down by Debye [9]. As is well known, in November of 1907, Einstein [10] proposed the first reasonably good model of the *Heat Capacity of a Solid* that employed the then nascent concept of quantization of energy. Einstein's [10] motivation was really not to propose a rigorous working model of a solid but to promote the then strange *Quanta Hypothesis* that had been promulgated earlier by Planck [1–3] and had been *given breath to* by him in his landmark and *1921 Nobel Prize* winning 1905 explanation of the *Photoelectric Effect* [11].

In his model of a solid, Einstein [10] made three fundamental assumptions: (1) Each atom in the lattice is an independent 3D quantum harmonic oscillator and the energy of this oscillator is quantized, (2) All atoms oscillate with the same fundamental frequency of vibration and (3) The probability of occupation of any given microstate is given by the Boltzmann thermodynamic probability. In summing up (integrating) all the energies of these oscillators, Einstein's oscillators have a minimum of zero frequency and an infinity frequency for a maximum frequency. While Einstein's [10] model gave a reasonably good fit to data, Debye [9] realized that Einstein's limits of integration where non-physical, especially the upper limiting frequency: $\nu_\mathrm{H} = \infty$. So, in constructing a revised (modified) version of Einstein's [10] model, Debye [9] had to correct this by limiting the upper frequency ν_H.

Debye [9] required that for the N oscillators – each with three degrees of freedom – the sum total of the modes of vibration must equal $3N$. That is to say, if $\mathrm{g}(\nu)$ is the density of states, then:

$$\int_{\nu_\mathrm{L}}^{\nu_\mathrm{H}} \mathrm{g}(\nu)d\nu = 3N. \tag{17}$$

Debye [9] set: $\nu_\mathrm{L} = 0$ because in reality: $\nu_\mathrm{L} \simeq 0$ and keeping ν_L as non-zero in his model did not bring in any significant improvement to the model, so he simply set this equal zero. Thus from (17), Debye [9] could calculate ν_H, and this maximum frequency one obtains from this calculation is known as the *Debye frequency* and symbolized ν_D.

For the photons under probe (in the present article), the density of states: $\mathrm{g}(\nu) = dN/d\nu$ can be calculated from (5)*, and so doing one obtains: $\mathrm{g}(\nu) = 8\pi V\nu^2/c^3$. Since a photon has two degrees of freedom, accordingly, N photons will have

---

*The reader must remember to substitute c in place of c_0 because in the foregoing calculation, we have disposed of the *vacuo* dispersion relation (1), and adopted the non-*vacuo* dispersion relation (2).

2N degrees of freedom, hence:

$$\int_{\nu_{\rm L}}^{\nu_{\rm H}} \frac{8\pi V \nu^2 d\nu}{c^3} = \frac{8\pi V}{c_0^3} \int_{\nu_{\rm L}}^{\nu_{\rm H}} n^3(\nu, \mathcal{T}) \nu^2 d\nu = 2N. \qquad (18)$$

Since $\nu_{\rm L}$ is known, $\nu_{\rm H}$ can be known if $n(\nu, \mathcal{T})$ is known. In the present article, we have no intention of evaluating the model, i.e. (16) and (18), that we have just set because we are yet to make further modifications where we shall include possible *non-zero photon mass effects*. For now, all we want to do is to show that one can demonstrate from a most fundamental level, that the emissivity function ϵ can be furnished with solid theoretical foundations rather than have this function as an experimental construct with no solid fundamental theoretical basis.

5 Discussion

The main aim of this paper has been to seek a fundamental and foundational basis and justification for the existence of the emissivity property of matter from the soils of fundamental theoretical physics. We are of the view that the grounds for such an endeavour have herein been set. Our final theoretically derived expression for the emissivity is given in (16). This expression we arrived at by revising the traditional derivation of the PRL as articulated in the introduction section. This emissivity function, i.e. (16), here derived has three free parameters associated with it and these parameters are:

1. The lower cut-off frequency: $\nu_{\rm L}$. The meaning of which is that there exists in this material medium in question, a *Lower Cutoff Frequency* ($\nu_{\rm L}$) below which frequency the body does not emit.
2. The upper frequency: $\nu_{\rm H}$. The meaning of which is that there exists in this material medium in question, an *Upper Cutoff Frequency* ($\nu_{\rm H}$) above which frequency the body does not emit.
3. The refractive index: n of the given material.

Of these three free adjustable parameters, the refractive index is less free as an adjustable parameter as there are already experimentally verified models of this quantity (see e.g. [12–14]). However, the lower ($\nu_{\rm L}$) and upper ($\nu_{\rm H}$) frequencies can be fixed to suit the given material, thus one can *in principle* fit the emissivity function (16) to the experimentally measured emissivity of a given material medium. When we say *one can in principle fit the emissivity function (16) to the experimentally measured emissivity of a given material medium*, we do not mean in an arbitrary manner, but that one will have to work out a realistic model that leads to a theory that fits to the data. In closing, allow us to say that in our next instalment, an attempt to fit the herein derived emissivity function, i.e. (16), to real data will be made.

6 Conclusion

Without the dictation of experience, it is possible in principle to justify by way of solid physical arguments and from a *bona fide* fundamental theoretic level, the existence and the need of the emissivity function for natural material.

Acknowledgements

We are grateful to the *Editorial Board* of the *Progress in Physics Journal* for the publication waiver on this article.

Received on August 6, 2020

References

1. Planck M. K. E. L. Entropie und Temperatur Strahlender Wärme. *Annalen der Physik*, 1900, v. 306 (4), 719–737.
2. Planck M. K. E. L. Ueber Irreversible Strahlungsvorgänge. *Annalen der Physik*, 1900, v. 306 (1), 69–122.
3. Planck M. K. E. L. Ueber das Gesetz der Energieverteilung im Normalspectrum. *Annalen der Physik*, 1901, v. 309 (3), 553–563.
4. Heisenberg W. K. Ueber den anschaulichen Inhalt der Quantentheoretischen Kinematik and Mechanik. *Zeitschrift für Physik*, 1927, v. 43 (3), 172–198. In English: Wheeler J. A. and Zurek W. H., eds. Quantum Theory and Measurement. Princeton University Press, Princeton, NJ, 1983, pp. 62–84.
5. Nyambuya G. G. Planck Radiation Formula for Massive Photons. *Prespacetime Journal*, 2017, v. 8 (5), 663–674.
6. Crisóstomo C. and Pitarma R. The Importance of Emissivity on Monitoring and Conservation of Wooden Structures Using Infrared Thermography. In: Advances in Structural Health Monitoring, IntechOpen, 2019, 16 pp. doi:10.5772/intechopen.82847.
7. Minissale M., Pardanaud C., Bisson R., and Gallais L. The Temperature Dependence of Optical Properties of Tungsten in the Visible and Near-Infrared Domains: An Experimental and Theoretical Study. *Journal of Physics D: Applied Physics*, 2017, v. 50 (45), 455601.
8. Sadiq H., Wong M. B., Tashan J., Al-MahaidiR. , and Zhao X.-L. Determination of Steel Emissivity for the Temperature Prediction of Structural Steel Members in Fire. *Journal of Materials in Civil Engineering*, 2013, v. 25 (2), 167–173.
9. Debye P. Zur Theorie der Spezifischen Wärmen. *Annalen der Physik*, 1912, v. 344 (14), 789–839.
10. Einstein A. Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Wärme. *Annalen der Physik*, 1907, v. 327 (1), 180–190.
11. Einstein A. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik*, 1905, v. 322 (6), 132–148.
12. Tang J., Sun M., Huang Y., Gou J., Zhang Y., Li G., Li Y., Man Y., and Yang j. Study on Optical Properties and Upconversion Luminescence of $\rm Er^{3+}/Yb^{3+}$ co-Doped Tellurite Glass for Highly Sensitive Temperature Measuring. *Optical Materials Express*, 2017, v. 7 (9), 3238.
13. Tan C. Z. and Arndt J. Temperature Dependence of Refractive Index of Glassy $\rm SiO_2$ in the Infrared Wavelength Range. *Journal of Physics and Chemistry of Solids*, 2000, v. 61 (8), 1315–1320.
14. Sellmeier W. Ueber die durch die Aetherschwingungen erregten Mitschwingungen der Körpertheilchen und deren Rückwirkung auf die ersteren, besonders zur Erklärung der Dispersion und ihrer Anomalien. *Annalen der Physik und Chemie*, 1872, v. 223 (11), 386–403.

Remark to Approach to the Schwarzschild Metric with SL(2,R) Group Decomposition

Alexander Kritov
E-mail: alex@kritov.ru

1 Remark to Section 5

1. The $SL(2,\mathbb{C})^*$ group definition. Let the group $SL(2,\mathbb{C})^*$ be a subgroup of $SL(2,\mathbb{C})$ with an element $Z' \in SL(2,\mathbb{C})^*$ such as

$$Z' = \left\{ \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} : a_1, a_4 \in Re,\ a_2, a_3 \in Im,\ \det(Z') = 1 \right\}.$$

The definition reflects the general Jacobian matrix form as given by (12) in [1].

2. The proof of the isomorphism of $SL(2,\mathbb{C})^*$ to $SL(2,\mathbb{R})$. The mapping (13) in [1] can be equivalently defined by the function that sends element of $Z' \in SL(2,\mathbb{C})^*$ to $Z \in SL(2,\mathbb{R})$

$$Z = T \cdot Z' \cdot T^{-1}$$

where

$$T = \begin{bmatrix} \sqrt{-i} & 0 \\ 0 & \sqrt{i} \end{bmatrix} \quad T^{-1} = \begin{bmatrix} \sqrt{i} & 0 \\ 0 & \sqrt{-i} \end{bmatrix} \quad \det(T) = 1.$$

The function is clearly a group homomorphism since

$$T \cdot Z'_1 \cdot Z'_2 \cdot T^{-1} = T \cdot Z'_1 \cdot T^{-1} \cdot T \cdot Z'_2 \cdot T^{-1} = Z_1 \cdot Z_2$$

for all $Z_1, Z_2 \in SL(2,\mathbb{R})$. It is obviously surjective. At last, as the inverse mapping

$$Z' = T^{-1} \cdot Z \cdot T$$

that sends any element of $SL(2,\mathbb{R})$ to $SL(2,\mathbb{C})^*$ is well defined it proves the injectivity. Hence, as a bijective homomorphism is shown, it finalizes the proof of $SL(2,\mathbb{C})^* \cong SL(2,\mathbb{R})$ mentioned in Section 5.

2 Corrections

The typo in the expression (10). The expression should evidently read with $\cosh^2(\beta)$ as follows

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

Section 5. A more appropriate notation for the Lorentz/Minkowski basis for $\mathbb{SL}(2,\mathbb{R})$ is $\mathbb{R}^{1(2)}$ as the group consists of the real numbers.

Received on Oct 11, 2020

References

1. Kritov A. Approach to the Schwarzschild Metric with SL(2,R) Group Decomposition. *Progress in Physics*, 2020, v. 16 (2), 139–142.

The Substantive Model of the Proton According to J. Wheeler's Geometrodynamic Concept

Anatoly V. Belyakov
E-mail: belyakov.lih@gmail.com

The review article presents the proton structure in accordance with the model based on a mechanistic interpretation of Wheeler's geometrodynamics. It is shown that this model gives a physically justified interpretation of the concepts introduced in quantum chromodynamics, such as "quark", "color", and also excludes the problem of confinement and others. The main parameters of the proton are calculated, namely: its mass, magnetic moment, lifetime, the proton-neutron mass difference, and also an analytical formula for its radius is derived. Typical lifetimes for various classes of elementary particles have been determined. Successful usage of the model gives reason to assume the model can be used for development a more rational theory of strong interactions instead of QCD.

1 Introduction

The internal structure of elementary particles, in particular of the proton, and their interactions with each other are considered in the theory of strong interactions (quantum chromodynamics). In QCD, there are no complete substantive models of the physical systems under consideration; instead, idealized virtual particles and quasiparticles (quarks, gluons) are introduced, as well as the concepts of an abstract image ("color") are appealed. QCD is based only on the observed properties of hadrons. It is assumed, that in hadrons interaction processes the numerous laws of conservation and redistribution (the amount of matter, energy, momentum, angular momentum, electric charge, magnetic flux and others) must simultaneously be fulfilled and various conditions must be observed.

In the absence of essential real models of elementary particles, the standard theory uses the approach of quantum mechanics: the properties of quarks and hadrons are simply described using wave functions and unitary symmetry combinatorics. The combination rules has formally been derived using the mathematical apparatus of quantum field theory and confirmed by experiments (there are 17 parameters that cannot be derived from the theory). However, it is not known why their physical nature is exactly this. In particular, it is not known why quarks can exist only in a bound state ("confinement"), which is recognized as one of the seven problems of the millennium.

This article shows the possibility of replacing the abstract QCD concepts applied to elementary particles with the particles real physical parameters. In contrast to the quantum theory, which states that micro-phenomena cannot be understood in any way from the point of view of our world scale, the mechanistic interpretation of Wheeler's idea first of all presupposes the presence of uniform or similar natural laws that are reproduced at different scale levels of matter. These laws, or at least their macroanalogues, are revealed in the structure of elementary particles. Therefore, there is reason to believe that the model based on Wheeler's idea can be used to construct a more rational theory of strong interactions instead of QCD.

2 On the macroanalogues

According to Wheeler's concept charges are considered as singular points on a three-dimensional surface, connected by a "wormhole" or vortex current tubes of the drain-source type, forming a generally closed contour, which a physical vacuum or some medium circulates along. From a purely mechanistic viewpoint the charge is proportional to the momentum of this medium in its motion along the vortex current tube contour, the spin, respectively, is proportional to the angular momentum of this medium relative to the contour longitudinal axis, while the magnetic interaction of the conductors is analogous to the forces acting between the current tubes.

The work [1] shows the possibility and expediency of introducing the "Coulombless" system of units and replacing the Coulomb with momentum. This approach allows using of well-known physical macroanalogues. The space and medium unit elements in the model are: an element with an electron mass m_e and size r_e and a vortex tube with a linear density $\varepsilon_0 = m_e/r_e$.

Microparticles are likened to vortex formations in an ideal liquid, where a vortex funnel (conditionally it is a surface X) is a *fermion analogue* with mass m_x, and a vortex thread in depth below the surface (conditionally it is a region Y) is a *boson analogue* with mass m_y, length l_y, radius r, and peripheral speed v. The vortex thread, in turn, is capable of twisting into a spiral forming subsequent structures (current tubes). In a real medium, these structures oscillate, transforming into each other (oscillation of oscillators); it is assumed that this is accompanied by the "packing" of the bosonic thread into a fermionic form. Apparently, fermion particles retain the bosonic part with half spin, which determines their magnetic and spin properties, and in the bosonic form the spin is restored to an integer value.

By the well-known physical analogy, the vortex tube of a contour, crossing over the surface of a liquid, creates ring waves or *contours of the next order*. Thus, interconnected contours are formed. Therefore, any particle seems to have two quantum numbers, depending on how one consider it: as a fermion (the analog of the proton being part of the greater contour of the subsequent family of particles) or as a boson (the mass of the contour of the previous family of particles). Thus, three generations of elementary particles as shown in [2] to form and there cannot be more. The microparticle itself is no longer considered as a point object and is characterized by the parameters of its own contour with a quantum number n.

The parameters of a bosonic vortex thread (or a contour with mass M) are determined in dimensionless units: in the fractions of the electron mass m_e, its classical radius r_e, and the speed of light c:

$$m_y = l_y = (an)^2, \tag{1}$$

$$v = \frac{c_0^{1/3}}{(an)^2}\,, \tag{2}$$

$$r = \frac{c_0^{2/3}}{(an)^4}\,, \tag{3}$$

where a is the inverse fine structure constant, c_0 is the dimensionless light velocity $c/[\mathrm{m/sec}]$.

In [3] a closed proton-electron contour is considered. From the condition of equality of the medium motion energy along the contour Mv^2 and the ultimate electron energy $m_e c^2$ the charge numerical value as the vortex current tube momentum and the projection angle value are determined. The projection angle value turned out to be complementary to the Weinberg angle $q_w \approx 28.7°$. Such a contour is "standard" and has parameters: the main quantum number $n = c_0^{1/3}/a = 4.884$, mass $M = c_0^{2/3} = 4.48 \times 10^5$, and the charge value (momentum)

$$e_0 = m_e c_0^{4/3} \cos q_w \times [\mathrm{m/sec}] = 1.602 \times 10^{-19}\ \mathrm{kg} \times \mathrm{m/sec}. \tag{4}$$

One can state therefore that the vortex current tube is formed by three vortex threads rotating around the common longitudinal axis. These threads are finite structures. They possess, by necessity, the right and left rotation; the last thread (it is evidently double one) possesses summary null rotation. These threads can be associated with vector bosons W^+, W^-, Z^0.

For the rotating unidirectional currents vortex threads with the condition of the magnetic and inertial (centrifugal) forces equilibrium their peripheral velocity v_0 is derived. If there are unit parameters, then it is true [2]:

$$v_0 = \frac{r_e}{(2\pi)^{1/2} \times [\mathrm{sec}]} = 1.124 \times 10^{-15}\ \mathrm{m/sec}. \tag{5}$$

This speed does not depend on the vortex threads length and on the distance between them. Thus, having some definite mass and length, bosonic vortex tubes *do not have a certain configuration and shape*. The latter indicates the difference between bosonic vortex tubes and their physical analogue; this is also the physical reason for their difference from fermions in that bosons do not obey the Pauli exclusion principle.

3 The proton and its parameters

With the extremely dense packing of a bosonic thread into a fermionic form, as shown in [2], the proton and electron own quantum numbers have the following values:

$$n_p = \left(\frac{2c_0}{a^5}\right)^{1/4} = 0.3338\,, \tag{6}$$

$$n_e = \left(\frac{2c_0}{a^5}\right)^{1/8} = 0.5777\,. \tag{7}$$

It was found that the relative mass of any fermion m_x with an arbitrary quantum number n_x is determined by the ratio:

$$m_x = \left(\frac{n_e}{n_x}\right)^{14}. \tag{8}$$

For a proton, as it turned out (with slight simplifications), its fermionic and boson masses are equal, $m_x = m_y = 2090$, which is the reason for its minimum baryon mass and its stability. When corrected by the Weinberg angle cosine, the proton relative mass is determined quite accurately, i.e. $m_p = 2090 \times \cos q_w = 1836$.

The charged particle included in a circulation contour is the place where a medium flow intersects the boundary between X- and Y-regions; there occurs a phase transformation. In this case, the fermionic and boson densities become equal, and the parameters of the medium acquire density and velocity *critical values*. The values of these critical parameters can be attributed to some quasiparticle — a *quark* that exists only in the phase transition region, which in fact is the part of the proton mass obtaining critical parameters. Moreover, in order to comply with the critical parameters at the standard contour energy $m_e c^2$, it is necessary to split the general contour current in the proton region into three parts (calculated value is 3.2). Under these conditions, the total quark mass m_k is 12.9. At the same time, as shown in [2] and [4], this mass depends on the interaction conditions and can take a minimum value equal to the electron mass.

In addition, the conditions for the flow continuity and charge constancy in any cross section of the contour (there must be three current lines) require the *reverse circulation currents* in the proton region to arise, which can be interpreted as the presence of zones with different charge signs in the proton. Using the minimal number of non-recurrent force current lines, one can schematically express current lines in a

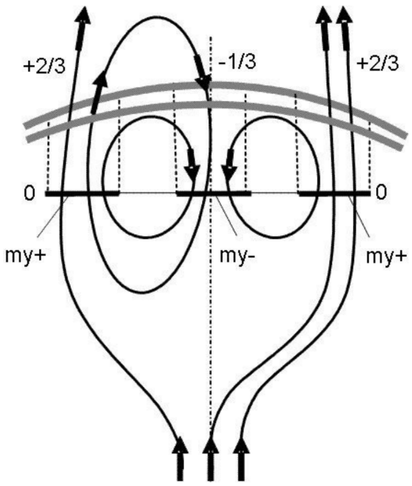


Fig. 1: A scheme of the proton: distribution of the current lines inside the proton.

proton in an unique way, as shown the Fig. 1. As seen, there exist two critical sections with a conditionally plus current (on the scheme up) and one section with a conditionally minus current (on the scheme down), where three current lines correspond to a general current in the scheme. Consequently, the proton fermionic surface for an external observer is as follows: the regions where force lines intersect the critical sections on the line 0–0 inside a proton will be projected onto the proton surface in the form of +2/3, +2/3, −1/3 of the total charge in according to the number and direction of the force lines crossing this surface. It would be more correct to associate quarks not with critical sections, but with stable ring currents containing, as follows from the diagram, one or two closed unit contours intersecting the critical sections.

Obviously, the proton parts in the critical sections have the velocity c and radius r_e, and they must have the bosonic vortex tubes of such length l_y that their own momentum would be equal to the momentum (charge) of an electron e_0. Assuming that the vortex tube linear density for critical conditions is proportional to the average quark mass $\frac{1}{3}\, m_k$, we can write:

$$\frac{1}{3}\, m_k\, \frac{m_e}{r_e}\, l_y r_e c = e_0\,. \tag{9}$$

Bearing in mind that in a "Coulombless" system the charge has the dimension of momentum and substituting the known parameter values, we find $l_y = 136.4$. Thus, the relative bosonic part length is actually equal to a, its unexcited mass is am_e, respectively, the vortex tube angular momentum (spin) is equal to $am_e c r_e = h/2\pi = \hbar$. Thus, the structures inside the proton are found with the relative length l_y and boson mass m_y, numerically equal to a, and the spin, equal to $\hbar$. Apparently, there are pairs of such boson tubes with the mass equal to that of a pion and with the counter-directional rotation that compensates for the spin. Depending on the currents impulse direction, they can form the pions family or be part of other mesons, which are supposed to exist in the close environment of protons in the form of a virtual meson "coat".

The *magnetic moment of the proton* μ_p in this model is calculated in accordance with its definition, where μ_p is the product (charge × velocity × path) and is determined by the bosonic configuration of the proton. The peripheral speed of the vortex threads relative to the Y-axis is v, the path is πr. Revealing v and r through (2) and (3), we finally get:

$$\mu_p = \frac{\pi c_0 e_0 c r_e}{(an_p)^6} = 1.39 \times 10^{-26}\ \mathrm{Am}^2. \tag{10}$$

For an electron, the path is the Bohr radius, and (10) takes the form:

$$\mu_e = \frac{\pi c_0 e_0 c R_B}{(an_e)^6} = 9.30 \times 10^{-24}\ \mathrm{Am}^2. \tag{11}$$

Only closed current lines remain in the neutron. The magnetic moment of the neutron equals two thirds of the proton's magnetic moment, i.e. proportional to the number of intersections of the critical sections by current lines (six instead of nine, existing in a proton, see Fig. 1) and is equal to -0.92×10^{-26} Am^2. Naturally, the magnetic moment sign changes in addition, because three positive open current lines are removed. The obtained values differ slightly from the actual ones, since the parameters n_p and n_e are determined with some simplifications.

The *neutron-proton mass difference* arises due to the acquisition of additional mass-energy by the neutron when the proton absorbs the electron. In [2] it was assumed that in the proton-neutron transition state one of the quark contours is located at the intersection of X–Y regions, and, becoming axisymmetric, increases itself to the maximum value m_{max}. In this case, keeping in mind (1) and (3), its parameters are $l_y = r = c_0^{2/9} = m_{max} = 76.5$. The difference between the kinetic energies of rotation of the excited contour and the initial quark contour with an average mass $\frac{1}{2}\, m_k$ should correspond (when corrected by the cosine of the Weinberg angle) to the proton-neutron mass difference Δm:

$$\left(m_{max} v_{max}^2 - \frac{1}{2}\, m_k v_k^2\right) \cos q_w = \Delta m. \tag{12}$$

Indeed, proceeding from their relative masses 76.5 and 12.9/2 and calculating their quantum numbers and velocities by (8) and (2), as a result, after substituting all quantities in (12), we find $\Delta m = 2.51$, which coincides with the actual value (2.53).

The exact *size of the proton* was determined in recent experiments [5]. It is significant that within the framework of this model, which does not use a complex mathematical apparatus and, in fact, is not a physical and mathematical model, but rather is a physical and logical one, it was possible to obtain an analytical formula for the proton size, proceeding from general laws only.

So, in [6] it was found that a single contour or a vortex tube having the momentum equivalent to the electron charge contains $n_i = 3$ single vortex threads, and the formula is obtained:

$$n_i = \frac{\left\{\left(m_e c_0^{2/3} r_e\right) / \left((2\pi)^{1/2} \times [\mathrm{sec}^2]\right)\right\}^{1/3}}{\{(2\pi)^{1/2} \gamma m_e^2 / r_e^2\}^{1/3}} = 2.973 \approx 3. \quad (13)$$

The formula is the cubic root of the ratio of the inertial forces, arising during the acceleration of the standard boson contour mass and acting towards the periphery (the value $r_e / \left((2\pi)^{1/2} \times [\mathrm{sec}]\right)$ is the rotation speed of the vortex threads relative to the contour longitudinal axis), to the gravity forces acting between masses of the size m_e at the distance r_e. The numerator is a constant value, so the formula depends only on the gravity forces, that is, on the interacting masses and the distance between them. It was shown in [4] that this ratio (or its modification for arbitrary m and r) can serve as a *coupling constant* equivalent, since it indicates the strength of bonds between the proton structure elements (quarks).

Moreover, it is the equality of these forces that determines the proton radius r_p and makes it possible to obtain an exact analytical formula under the condition $m_k = m_e$. Let us specify formula (13) under the assumption that the quarks are located at the corners of a regular triangle and that each of them is affected by the sum of two projections of forces, and also take into account that r_p is the size of the circumscribed circle. Then the formula is written in the form of equality of the dynamic and gravitational forces:

$$\frac{m_e c_0^{2/3} r_e}{2\pi \times [\mathrm{sec}^2]} = \frac{2 \sin 60^\circ \gamma m_e^2}{(r_p \sin 60^\circ)^2}\,, \quad (14)$$

whence we get:

$$r_p = \frac{(8\pi\gamma\varepsilon_0)^{1/2} \times [\mathrm{sec}]}{3^{1/4} c_0^{1/3}} = 0.836 \times 10^{-15}\ \mathrm{m}, \quad (15)$$

which exactly coincides with the proton charge radius value obtained in the recent experiments (0.833 femtometers, with an uncertainty of ±0.010 femtometers [5]). Thus, the boson mass of the standard contour, which is in the Y region and is, as it were, hidden, having the value $m_e c_0^{2/3}$, determines not only the proton charge and spin, but also its radius. Note that this radius is determined for the proton in the hydrogen composition, but not for a single proton, where it can have a greater value. It is important that the formula (15) contains the gravitational constant; in papers [2, 7, 8] the necessity of introducing gravity into the microcosm is shown, in particular, to determine the neutrino mass.

Thus, there is a sufficient set of parameters for the proton internal structure to describe the strong interaction. The concepts of fractional charge, quarks and color find here their physical representation. Indeed, there are two different ring currents or circuits (quarks u and d), the force lines of which are projected onto the proton outer surface in the form of fractional charges, and three different critical sections ("colored" quarks). Moreover, as the contour currents can be directed in the opposite direction, forming antiquarks, so the vortex tubes in the critical sections can have the opposite direction of rotation, creating an "anticolor". It is obvious that the proposed proton structure in the form of a field lines unique configuration no longer requires the confinement existence and, consequently, the filling of the proton region with the "sea" of virtual quarks and gluons.

In fact, the concept originating from the hydrodynamics of a continuous inviscid media is proposed here and this analogy turned out to be correct. Moreover, it has been established that the light velocity can be calculated by the equation describing the wave propagation on a liquid surface [9].

4 On the elementary particles lifetime

The microparticles decay probability and their lifetimes depend on many factors. The most important of them is the type of interaction (electromagnetic, weak, strong), which is responsible for the decay that occurs. The lifetimes of elementary particles differ extremely strongly: $10^{-6} \dots 10^{-25}$ seconds, at that most of them are grouped according to their lifetimes in rather narrow intervals. This model has objective parameters that allow one to estimate the microparticle various classes lifetime. Further there are calculated values, in general, corresponding to the average lifetimes for these classes.

The *microparticle lifetime* t (except for resonances and W, Z bosons) can be estimated as the time it takes to run around with a velocity v over the entire "stretched" contour length [7]:

$$t = \frac{a^8 n^8}{c_0} \frac{r_e}{c}\,. \quad (16)$$

But W, Z bosons and resonances decay even before the final spiral structure is formed, i.e. they are, as it were, not completely particles. W, Z bosons have the shortest decay time, and it is determined by the time it takes to run with the speed of light around the electron vortex tube with the radius r. Bearing in mind (3),

$$t_{min} = \pi \frac{r_e}{c} \frac{c_0^{2/3}}{(a n_e)^4} = 3.4 \times 10^{-25}\ \mathrm{sec}. \quad (17)$$

For numerous resonances the lifetime correlates well with the run time with the light velocity of the contour radius $l_y / 2\pi$. Since $l_y = m_y$, then

$$t = \frac{m_y r_e}{2\pi c}\,. \quad (18)$$

For example, for Y, J/Ψ, η-particles with masses $m_y = 19700, 6056, 1074$ values $t = 2.95 \times 10^{-20}$, 0.91×10^{-20}, 0.30×10^{-20} seconds.

In the group of heavy hadrons, particles contain unstable heavy quarks, and they decay through rapid weak decays. Then, in formula (16) for a weak decay n must be minimal, i.e. equal to 1.643 [7], and $t = 2.1 \times 10^{-13}$ seconds.

Light and "strange" hadrons are more stable, and in formula (16) the parameter n should have the value of its own Y-contour [4]. For a group, on average, $n \approx 3.5$, and $t \approx 10^{-10}$ seconds.

Particles that decay due to strong interaction, for example η and π^0-particles, live only within the proton or electron own contours. Therefore, for them, when substituting the values of n_p and n_e in (16), the values of t is about 6×10^{-19} and 5×10^{-17} seconds.

Finally, during the electromagnetic decay of light charged particles (pions, kaons) the contour with large n and, accordingly, with the largest value of t manages to form.

As for the *neutron lifetime*, it is assumed [2] that the neutron loses the acquired mass-energy Δmc^2 gradually with fractions of $m_e v^2 c^2$ for a time per each fraction equal to the vortex threads rotation period inside the current tube r_e/v_0. Bearing in mind (2) and (5), we obtain the duration of the total energy dissipation by the neutron:

$$t = (2\pi)^{1/2} \, \frac{\Delta m}{\cos q_w} \, \frac{a^4 n_e^4}{c_0^{2/3}} \times [\text{sec}] = 629 \text{ sec.} \tag{19}$$

The same duration is determined by the time constant - the return duration of the excited axisymmetric contour with the total length $\pi c_0^{2/9}$ to its initial state due to its constituent current lines rotation with the speed v_0:

$$t = \frac{\pi c_0^{2/9} r_e}{v_0} = 604 \text{ sec.} \tag{20}$$

The neutron half-life is about 609 seconds. Thus, the consistency of formulas (12), (19), (20) with each other and their results with the actual values of the neutron lifetime and the neutron-proton mass difference confirm the accepted model of proton-neutron transitions.

5 Conclusion

In the articles concerning the microworld and, in particular, the proton properties, it has been established that there are only three generations of elementary particles. The parameters of the proton (mass, magnetic moment, charge radius, proton lifetime, neutron-proton mass difference) are determined. A physical explanation is proposed for such abstract images as quarks and their confinement, "color", pions "coat", etc. The results were obtained within the framework of the elementary model based on the mechanistic interpretation of Wheeler's geometrodynamics. Wherein the balances between the main interactions and general patterns were only used, moreover, without adding any empirical coefficients.

The model can be used as a basis for constructing the theory of strong interactions, which can be an alternative for QCD. In a possible new theory, the interpretation of the concepts and results obtained, which form the model basis, can be performed in some terms of electrodynamics (or some other) on the basis of a suitable mathematical apparatus.

Submitted on November 4, 2020

References

1. Belyakov A.V. On the uniform dimension system. Is there the necessity for Coulomb? *Progress in Physics*, 2013, v.9, issue 3, 142–143.
2. Belyakov A.V. Macro-analogies and gravitation in the micro-world: further elaboration of Wheeler's model of geometrodynamics. *Progress in Physics*, 2012, v.8, issue 2, 47–57.
3. Belyakov A.V. Charge of the electron, and the constants of radiation according to J. A. Wheeler's geometrodynamic model. *Progress in Physics*, 2010, v.6, issue 4, 90–94.
4. Belyakov A.V. Nuclear power and the structure of a nucleus according to J. Wheeler's geometrodynamic concept. *Progress in Physics*, 2015, v.11, issue 1, 89–98.
5. Bezginov N., Valdez T., Horbatsch M., Marsman A., Vutha A.C., Hessels E.A. A measurement of the atomic hydrogen Lamb shift and the proton charge radius. *Science*, 06 Sep. 2019, v.365, issue 6457, 1007–1012.
6. Belyakov A.V. On materiality and dimensionality of the space. Is there some unit of the field? *Progress in Physics*, 2014, v.10, issue 4, 203–206.
7. Belyakov A.V. Gravity in the microworld. *Progress in Physics*, 2020, v.16, issue 1, 58–61.
8. Belyakov A.V. Determination of the neutrino mass. *Progress in Physics*, 2016, v.12, issue 1, 34–38.
9. Belyakov A.V. On the speed of light and the continuity of physical vacuum. *Progress in Physics*, 2018, v.14, issue 4, 211–212.

Algebra of Discrete Symmetries in the Extended Poincaré Group

Valeriy V. Dvoeglazov

UAF, Universidad Autónoma de Zacatecas, Apartado Postal 636, Zacatecas 98061 Zac., México
E-mail: valeri@fisica.uaz.edu.mx

We begin with the comprehensive review of the basics of the Lorentz, (extended) Poincaré Groups and O(3,2) and O(4,1). On the basis of the Gelfand-Tsetlin-Sokolik-Silagadze research [1-3], we investigate the definitions of the discrete symmetry operators both on the classical level, and in the secondary-quantization scheme. We study physical content within several bases: light-front form formulation, helicity basis, angular momentum basis, on several practical examples. The conclusion is that we have ambiguities in the definitions of the corresponding operators P, C; T, which lead to different physical consequences.

Talk presented at the LXII Congreso Nacional de Física. 6–11/10/2019. Villahermosa, Tab., México.

1 The standard definitions

The Lorentz Group conserves the interval $ds^2 = dx^\mu dx_\mu$ in the 4-space with respect to (pseudo) Euclidean rotations. The Poincaré Group includes translations in the Minkowski space. The extended Poincaré Group includes discrete transformations, the unitary C, P, and the antiunitary T in quantum field theory (QFT). The P is the space inversion: $x^0 \to x^0$, $\mathbf{x} \to -\mathbf{x}$. The T is the time reversal: $x^0 \to -x^0$, $\mathbf{x} \to \mathbf{x}$. The C is the electric charge conjugation. It is related to the PT operation: $x^0 \to -x^0$, $\mathbf{x} \to -\mathbf{x}$. The interval is also conserved under these operations. In QFT, the eigenvalues of the combined CPT are also invariants.

While [4] presented the derivation method to obtain the field operator *ab initio*, we define the field operator [5, 6] in the pseudo-Euclidean metrics as follows:

$$\Psi(x) = \sum_h \int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_p} \left[u_h(\mathbf{p}) a_h(\mathbf{p}) e^{-ip\cdot x} + v_h(\mathbf{p}) b_h^\dagger(\mathbf{p}) e^{+ip\cdot x}\right] . \tag{1}$$

Hence, the Dirac equation is:

$$\left[i\gamma^\mu \partial_\mu - m\right] \Psi(x) = 0 . \tag{2}$$

At least, 3 methods of its derivation exist [7–9]:

- the Dirac method (the Hamiltonian should be linear in $\partial/\partial x^i$, and be compatible with $E_p^2 - \mathbf{p}^2 c^2 = m^2 c^4$);
- the Sakurai one (based on the equation $(E_p - \sigma\cdot\mathbf{p})(E_p + \sigma\cdot\mathbf{p})\,\phi = m^2\phi$);
- the Ryder one (the relation between 2-spinors at rest is $\phi_R(\mathbf{0}) = \pm\phi_L(\mathbf{0})$ and boosts).

It has solutions of positive energies and negative energies. The latter are reinterpreted as the antiparticles.

$$E_p = \sqrt{\mathbf{p}^2 + m^2},\ c = \hbar = 1,\ g^{\mu\nu} = \mathrm{diag}\{1, -1, -1, -1\} .$$

The solutions in the momentum representation are: $u_h(\mathbf{p}) = \mathrm{column}(\phi_R^h(\mathbf{p}), \phi_L^h(\mathbf{p}))$. Next,

$$u_h = \begin{pmatrix} \exp(+\boldsymbol{\sigma}\cdot\boldsymbol{\varphi})\,\phi_R^h(\mathbf{0}) \\ \exp(-\boldsymbol{\sigma}\cdot\boldsymbol{\varphi})\,\phi_L^h(\mathbf{0}) \end{pmatrix},\ v_h(\mathbf{p}) = \gamma^5 u_h(\mathbf{p}) , \tag{3}$$

where $\cosh(\varphi) = E_p/m$, $\sinh(\varphi) = |\mathbf{p}|/m$, $\hat{\varphi} = \mathbf{p}/|\mathbf{p}|$, and h is the polarization index. It is shown that the parity operator can be chosen as

$$P = e^{i\alpha_s}\gamma^0 R,\ \gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \tag{4}$$

because

$$\left[i\gamma^\mu \partial'_\mu - m\right] \Psi(x^{\mu'}) = 0 , \text{(change of variables)} , \tag{5}$$

where

$$\Psi(x^{\mu'}) = \mathcal{A}\,\Psi(x^\mu) , \text{(linearity)} . \tag{6}$$

These conditions can be satisfied by the γ^0 matrix in the Weyl basis. R can be chosen

$$R \equiv (\theta \to \pi - \theta, \phi \to \pi + \phi, r \to r) .$$

For fermions, it is well known that a particle and an antiparticle have opposite eigenvalues of the parity operator in this $(1/2, 0) \oplus (0, 1/2)$ representation of the Lorentz Group. In QFT we should have:

$$U_P \psi(x) U_P^\dagger = e^{i\alpha_s}\gamma^0 \psi(x') . \tag{7}$$

So,

$$\begin{aligned} U_P a_h(\mathbf{p}) U_P^\dagger &= e^{+i\alpha_s} a_h(\mathbf{p}') , \\ U_P b_h(\mathbf{p}) U_P^\dagger &= -e^{-i\alpha_s} b_h(\mathbf{p}') . \end{aligned} \tag{8}$$

The operator U_P can be constructed in the usual way, see [5] and [6]. The charge operator interchanges the particle and the antiparticle. For example, in the Dirac case on the classical level:

$$u_\uparrow \to -v_\downarrow,\ u_\downarrow \to +v_\uparrow , v_\uparrow \to +u_\downarrow,\ v_\downarrow \to -u_\uparrow . \tag{9}$$

Thus, we can write, thanks to E. Wigner:

$$C_{1/2} = e^{i\alpha_c}\begin{pmatrix}0 & i\Theta\\ -i\Theta & 0\end{pmatrix}\mathcal{K}, \quad \Theta = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} = -i\sigma_2\,. \tag{10}$$

In QFT, we should have:

$$U_C \psi(x) U_C^\dagger = e^{i\alpha_c} C \psi^\dagger(x)\,. \tag{11}$$

So [5],

$$U_C a_h(\mathbf{p}) U_C^\dagger = e^{+i\alpha_c} b_h(\mathbf{p})\,,\ U_C b_h(\mathbf{p}) U_C^\dagger = e^{-i\alpha_c} a_h(\mathbf{p})\,. \tag{12}$$

See however [11], where two possibilities for the charge conjugation operator have been proposed.

The time reversal operator is antiunitary (see Wigner and [4]). Let us remind that the operator of hermitian conjugation does not act on c-numbers on the left side of (13) below. This fact is connected with the properties of an antiunitary operator: $\left[V^T \lambda A (V^T)^{-1}\right]^\dagger = \left[\lambda^* V^T A (V^T)^{-1}\right]^\dagger = \lambda\left[V^T A^\dagger (V^T)^{-1}\right]$.

$$\left[V^T_{[1/2]}\Psi(x^\mu)(V^T_{[1/2]})^{-1}\right]^\dagger = S(T)\,\Psi^\dagger(x^{\prime\prime\mu})\ . \tag{13}$$

We can see that C and P anticommute in the Dirac case:

$$\{C, P\}_+ = 0\,, P^2 = 1\,, C^2 = 1\,, \tag{14}$$

and $(CPT) = \pm 1$. However, we present the opposite case later, where $(CPT) = \pm i$, which is related to the commutation (anticommutation) of the C and P operators.

The table on p. 157 of [5] gives us the properties of the scalar, 4-vector, tensor, axial-vector and pseudoscalar under these transformations in the case of the "Dirac-like parity" definitions. However, see the next Section.

2 Anomalous representations of the inversion group

The previous Section perfectly describes the CPT properties of the charged fermions. Nevertheless, the authors of [1,2,10] proposed another class of representations of the full Lorentz Group long ago. As it was shown recently, it may be applied to the (anti)bosons of the opposite parities, and to the (anti)fermions of undefined parities. The latter are not the eigenstates of the parity operator, but they are the eigenstates of the charge-conjugate operator. Gelfand, Tsetlin and Sokolik noted that there exist representations of the full Lorentz Group of the anomalous parity. Originally, this concept was intended to be applied to explain the decay of $K-$mesons.

The examples are: one can note that in the $(1/2, 1/2)$ representation (or for x^μ) the operators of the space inversion (t_{01}), the time reversal (t_{10}) and the combined space-time inversion (t_{11}) are commutative. They form the inversion group together with the unit element. Let us search the projective representations of the Lorentz group combined with the discrete group. As opposed to the usual case, $t_{01}t_{10} = t_{11}$, $t_{10}t_{11} = t_{01}$, $t_{01}t_{11} = t_{10}$, for instance, one can drop the condition of commutativity, and one can form the projective representation with $T_{10}T_{01} = -T_{11}$, or $T_{11}T_{11} = -1$, see the full table in [1]. They noted that there are *two* normal-parity (in their notation) and *two* anomalous parity representations for (bi)spinors. Then, they extended the concept of the anomalous parity to any representation of the proper Lorentz Group characterized by the numbers (k_0, k_1) and $(-k_0, k_1)^*$. When

$$[T_{i'k'}, T_{i''k''}]_+ = 0\,, \tag{15}$$

this is the case of the anomalous parity (later, this was confirmed by Nigam and Foldy [12]). G. Sokolik noted that this concept is related to the concept of the 5-D representations of the proper orthogonal group with pseudo-Euclidean metrics. For example,

$$\begin{aligned} &T_{10} \sim H_{54} = \exp(i\pi I_{54}/2)\,,\\ &T_{11} \sim H_{43}H_{21} = \exp(i\pi I_{43})\exp(i\pi I_{21})\,,\\ &T_{01} = T_{11}T_{10}\,. \end{aligned} \tag{16}$$

T_{10}, T_{01}, T_{11} leave invariant the extended 8-component Dirac equation only (compare with [13] and [14]):

$$\Gamma_\mu \partial^\mu \psi + m\psi = 0\,,\ \Gamma_\mu = \begin{pmatrix}\gamma_\mu & 0\\ 0 & -\gamma_\mu\end{pmatrix}. \tag{17}$$

They claimed relations to the concepts (known in that time):

- istopic spin;
- fusion theory;
- the non-linear Heisenberg theory

were mentioned. The corresponding matrix representations of the anomalous-parity representations have been presented:

$$T_{01} = \begin{pmatrix}0 & I\\ I & 0\end{pmatrix},\ T_{10} = \begin{pmatrix}0 & -I\\ I & 0\end{pmatrix},\ T_{11} = \begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}, \tag{18}$$

and

$$T_{01} = \begin{pmatrix}0 & -iI\\ iI & 0\end{pmatrix},\ T_{10} = \begin{pmatrix}0 & iI\\ iI & 0\end{pmatrix},\ T_{11} = \begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}. \tag{19}$$

Later Wigner [10] repeated their results in the Istanbul School lectures (1962), and Silagadze [3] rediscovered and applied this possibility in 1992. The conclusion of these papers is: we noted that both new versions of the representations of the full Lorentz Group (commuting spinor and anticommuting boson representations) lead to the doubling of the dimensionality of the $\psi-$function.

3 The self/anti-self charge conjugate states

The content of this Section contains the material of [11]. The conclusions are: we have constructed another explicit example of the BWW-GTS theory. The matter of physical dynamics connected with this mathematical construct should be solved in the future, dependent on what gauge interactions with potential fields we introduce [14] and what experimental setup we choose.

Received on Nov. 6, 2020

References

1. Gelfand I. M. and Tsetlin M. L. *Sov. Phys. JETP*, 1957, v. 4, 947.
2. Sokolik G. A. *Sov. Phys. JETP*, 1958, v. 6, 1170; ibid., 1959, v. 9, 781; *Dokl. Akad. Nauk SSSR*, 1957, v. 114, 1206.
3. Silagadze Z. K. *Sov. J. Nucl. Phys.*, 1992, v. 55, 392.
4. Bogoliubov N. N. and Shirkov D. V. Introduction to the Theory of Quantized Fields. John Wiley & Sons, NY, 1980.
5. Itzykson C. and Zuber J.-B. Quantum Field Theory. McGraw-Hill International Book Co., NY, 1980.
6. Greiner W. Field Quantization. Springer, Berlin-Heidelberg, 1996.
7. Dirac P. A. M. *Proc. Roy. Soc. Lond.*, 1929, v. A117, 610.
8. Sakurai J. J. Advanced Quantum Mechanics. Addison-Wesley, Reading, 1967, §3.2.
9. Ryder L. H. Quantum Field Theory. Cambridge University Press, Cambridge, 1985.
10. Wigner E. P. in: Gursey F., ed., Group Theoretical Concepts and Methods in Elementary Particle Physics. Gordon and Breach, NY, 1964.
11. Dvoeglazov V. V. *Mod. Phys. Lett.*, 1997, v. A12, 2741. arXiv: hep-th/9609142.
12. Nigam B. P. and Foldy L. L. *Phys. Rev.*, 1956, v. 102, 1410.
13. Markov M. A. *ZhETF*, 1937, v. 7, 579; ibid., 1937, v. 7, 603.
14. Dvoeglazov V. V. *Nuovo Cim.*, 1995, v. 108A, 1467.

Theoretical Study on Polarized Photon

Shixing Weng

E-mail: wengs2015@gmail.com, Brampton, Canada

A solution of electromagnetic four-potential for polarized photon is obtained by solving its wave equations in elliptic cylindrical coordinates. An explicit energy wave function for the photon is presented in the form of a linear combination of the electric field and magnetic field from the solution. This wave function is used to calculate the angular momentum value of the photon. The elliptic coordinate parameter, a, for the photon is considered to be equal to a quarter of its wavelength.

1 Introduction

Photon as a quantum of light has attracted many researchers to develop explanations on its behaviors and to experiment to determine its properties. The photon as a fundamental wave-particle which moves at the speed of light serves like a messenger traveling from one place to another, which is necessary for the physical world to work properly. The classical view on light is provided by Maxwell's theory of electromagnetism [1], hence light is considered as a bundle of electromagnetic transverse waves. The particle view of light in modern physics may be provided by Einstein [2], so a photon has not only energy but also momentum. Work has been done to unify these two views. An expression for photon wave function is introduced by using the Riemann-Silberstein vector which is a linear combination of the electric field and magnetic field of the photon. An overview of the work on photon wave function is available in [3].

A photon has wave-particle duality which may be explained by a single entity as a joint wave-particle [4]. A more specific view on the electromagnetic structure for the photon is presented in [5], which is for circularly polarized photons. Hence the photon in circular polarization may be viewed as a charged moving electric capacitor with electric charge distributed circularly on its cylindrical surface of radius $\lambda/2\pi$, where λ is the wavelength of the photon.

In this article, we present our theoretical study on polarized photons. It is well known that polarized light has the property of certain orientation which may be generated by an optical polarizer. Recent experiment [6] shows that the transmission intensity of polarized light strongly correlates to the orientation of elliptically-shaped holes on the transmission plate. This as an example indicates that the transverse field strength of photons in the polarized light is not circularly distributed evenly as different from that of circularly polarized photons. The novelty of this article is on: the wave equation for the photon is solved within the elliptic cylindrical coordinates; an explicit photon energy wave function is presented based on the expression of Riemann-Silberstein vector wave function (in the next section); quantum expressions of the energy density, energy current density and the angular momentum or spin density for the photon are derived from the wave function. We are not aware of such work in the literature.

This article is divided into the following sections: Introduction, Method, Results and Discussions, and Conclusion. The Introduction section provides a brief overview on our current understanding of the photon.

In the Method section, we will use similar method as in [5]. First we obtain a solution for the electromagnetic four-potential by solving the wave equations in elliptic cylindrical coordinates. The electromagnetic four-potential generally includes a scalar potential, which is an electric potential divided by the speed of light, and a vector potential. Then show to get the electric field and magnetic field from the solution of the four-potential; an explicit energy wave function for the photon is presented as a linear combination of the electric field and magnetic field; other expressions such as photon energy density, energy current density and angular momentum density are derived based on quantum mechanics.

In the Results and Discussions section we show the results for the photon expressions developed in the previous section, such as the four-potential, electromagnetic fields, the wave function, energy and energy current densities, and angular momentum for the photon; fairly detailed work is presented in evaluating the angular momentum value for the photon; some particularities are discussed. The Conclusion section provides a brief summary of the work presented in this article. We use MKS units in this work.

2 Method

In the space region where there are no other free electric charge and electric current, the electric potential ψ and the vector potential $\mathbf{A}$ satisfy the following wave equations, respectively,

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \nabla^2\psi = 0\,, \tag{1}$$

$$\frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} - \nabla^2\mathbf{A} = 0\,, \tag{2}$$

where c is the speed of light, t is time, ∇^2 is the Laplacian operator, and $\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$ is D'Alembert's operator which is also written as $\square$. In obtaining these equations the set of Maxwell equations with Lorenz gauge is employed. The Lorenz gauge

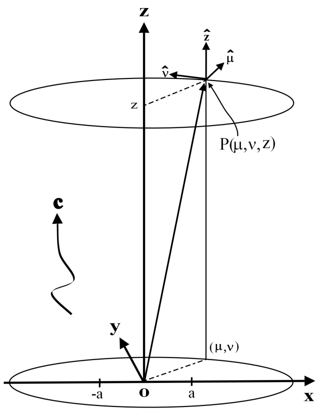


Fig. 1: A drawing of the elliptic cylindrical coordinate system together with the cartesian coordinates, where $\hat{\mu}$, $\hat{\nu}$, and $\hat{\mathbf{z}}$ are unit vectors for the coordinate system and a is a length parameter that marks the focal points on $\mathbf{x}$ of the ellipse. The major axis of the ellipse is $\mathbf{x}$. The wave symbol represents a photon moving in the direction of the positive $\mathbf{z}$ axis at the speed of light c.

is given by

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \psi}{\partial t} = 0\,. \tag{3}$$

Eqs. (1) and (2) are satisfied with solutions for traveling waves.

For the polarized photon, we solve (1) and (2) in elliptic cylindrical coordinates as shown in Fig. 1. Where the relationships between the cartesian and elliptic cylindrical coordinates are

$$\begin{aligned} x &= a\cosh\mu\cos\nu\,, \\ y &= a\sinh\mu\sin\nu\,, \\ z &= z\,, \end{aligned} \tag{4}$$

where x, y, z are cartesian coordinate values and $\mu\,\nu, z$ are elliptic cylindrical coordinate values, a is a length parameter which specifies the focal points of the ellipse, $\mu \in (0,\infty)$ and $\nu \in (0, 2\pi)$. The value of a will be considered later to be proportional to the wavelength of the photon. The scale factors are

$$\begin{aligned} h_\mu &= h_\nu = a\gamma\,, \\ h_z &= 1\,, \end{aligned} \tag{5}$$

where $\gamma = \sqrt{\sinh^2\mu + \sin^2\nu}$.

We find for this particular case that the vector potential $\mathbf{A}$ has a z component only so $\mathbf{A} = \hat{\mathbf{z}}A_z$ and $\nabla^2\mathbf{A} = \hat{\mathbf{z}}\nabla^2 A_z$, where $\hat{\mathbf{z}}$ is the unit vector for the z axis. The Laplacian operator ∇^2 for the elliptic cylindrical coordinates is expressed as

$$\nabla^2 = \frac{1}{a^2\gamma^2}\left(\frac{\partial^2}{\partial\mu^2} + \frac{\partial^2}{\partial\nu^2}\right) + \frac{\partial^2}{\partial z^2}\,. \tag{6}$$

Hence (1) and (2) in the elliptic cylindrical coordinates are satisfied with the following general solution:

$$f = f_0\, e^{-\mu}\sin(\phi)\,, \tag{7}$$

where f is a general quantity that may represent either ψ or A_z here, f_0 is the corresponding constant, $\phi = kz + \nu - \omega t$, and $k = \omega/c$, and ω is the angular frequency of the photon. We choose the "–" sign in the exponential function to make the solution to be limited in space. Here we let the photon travel in the z direction. And we arbitrarily choose the sine function here, one may choose cosine function as well but the results should be similar. By using the Lorenz gauge we have the following relationship for the electric potential constant, ψ_0, and the vector potential constant, A_0, as

$$A_0 = \psi_0/c\,. \tag{8}$$

Once we have the solution of the four-potential we can calculate [7] the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ using the following equations,

$$\mathbf{E} = -\nabla\psi - \frac{\partial\mathbf{A}}{\partial t} = -\frac{1}{a\gamma}\left(\hat{\mu}\frac{\partial}{\partial\mu} + \hat{\nu}\frac{\partial}{\partial\nu}\right)\psi\,, \tag{9}$$

$$\mathbf{B} = \nabla\times\mathbf{A} = \frac{1}{a\gamma}\left(\hat{\mu}\frac{\partial}{\partial\nu} - \hat{\nu}\frac{\partial}{\partial\mu}\right)A_z\,, \tag{10}$$

where $\hat{\mu}, \hat{\nu}$ are unit vectors for μ and ν, respectively, and "×" represents the vector cross operator. In deriving (9) for the electric field, we used this case relationship:

$$\frac{\partial\psi}{\partial z} + \frac{\partial A_z}{\partial t} = 0.$$

Both the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ are vectors with μ and ν components, which are perpendicular to the direction of the wave propagation. They represent transverse waves.

As we know, a photon is a packet of energy in electromagnetic field form and moves at the speed of light. This means that the electric field $\mathbf{E}$ or the magnetic field $\mathbf{B}$ of the photon can not exist alone and they are both like two faces of one body. We have the following expression of the electromagnetic field $\mathbf{F}$ suit for the photon

$$\mathbf{F} = \frac{1}{\sqrt{2}}\left(\sqrt{\epsilon}\mathbf{E} + i\frac{\mathbf{B}}{\sqrt{u}}\right), \tag{11}$$

where ϵ is the permittivity and u is the permeability in the space region where photon absorption is negligible, and i is the imaginary unit. This expression is known as the Riemann-Silbertein vector and was introduced as a photon wave function in [8]. Here the choice of "+" sign for the imaginary part is arbitrary, one may choose "−" for similar results. Like **E** or **B**, **F** is also a vector which satisfies the wave equation and also represents a transverse traveling wave. The field **F** is a complex vector in general and is characterized as a quantum vector wave function. Hence methods developed in quantum mechanics may be employed here [9]. By the dimensional analysis we know that **F** represents an energy density wave function. In the following we use **F** to derive expressions for energy and current densities and then the angular momentum for the photon. For clarity, the cartesian coordinates are used in the following work. We start from the wave equations:

$$\frac{1}{c^2}\frac{\partial^2 \mathbf{F}}{\partial t^2} - \nabla^2 \mathbf{F} = 0\,, \tag{12}$$

and

$$\frac{1}{c^2}\frac{\partial^2 \mathbf{F}^*}{\partial t^2} - \nabla^2 \mathbf{F}^* = 0\,, \tag{13}$$

where $\mathbf{F}^*$ is the conjugate of **F**. And

$$\mathbf{F} = \hat{\mathbf{x}} F_x + \hat{\mathbf{y}} F_y\,, \tag{14}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are unit vectors and F_x, F_y are the field components for x and y axes, respectively. As a transverse wave, **F** has x and y components only and the z component, F_z is zero. Since our original solution for **F** is in elliptic coordinates with components of μ and ν, we may convert those to x and y components using the following matrix multiplication,

$$\begin{bmatrix} \hat{\mu} \\ \hat{\nu} \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} \sinh\mu\cos\nu & \cosh\mu\sin\nu \\ -\cosh\mu\sin\nu & \sinh\mu\cos\nu \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \end{bmatrix}. \tag{15}$$

Since F_x and F_y are explicit functions of μ and ν, in order to do their derivatives with respect to x and y we need partial derivatives of μ and ν to x and y by using the following matrix form:

$$\begin{bmatrix} \delta\mu \\ \delta\nu \end{bmatrix} = \frac{1}{a\gamma^2} \begin{bmatrix} \sinh\mu\cos\nu & \cosh\mu\sin\nu \\ -\cosh\mu\sin\nu & \sinh\mu\cos\nu \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}, \tag{16}$$

where δ is a tiny increment. In obtaining (16), we first do the tiny variations of (4) for x and y to μ and ν to get a conversion matrix between the two coordinate systems. And then find the inverse matrix as in (16). Eq. (15) is equivalent to (16) if we replace each variation together with its scale factor such as $a\gamma$ in the latter equation by the corresponding unit vector.

As is common in quantum mechanics to find the energy density and the energy current density for the photon, we do this operation:

$$\mathbf{F}^* \cdot (12) - \mathbf{F} \cdot (13)\,, \tag{17}$$

where "·" represents the dot product operator and "*" is the complex conjugate symbol, and

$$\mathbf{F}^* \cdot \nabla^2 \mathbf{F} = \sum_i^{x,y,z} F_i^* \, \nabla^2 F_i\,, \tag{18}$$

where the summation is over the three cartesian components. By a few mathematical operations, we have the following form of energy current and density continuity equation:

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \tag{19}$$

with

$$\mathbf{j} = \frac{c^2}{2i\omega} \sum_i^{x,y,z} \left(F_i \, \nabla F_i^* - F_i^* \, \nabla F_i \right) \tag{20}$$

and

$$\rho = \frac{1}{2i\omega} \left(\mathbf{F}^* \cdot \frac{\partial \mathbf{F}}{\partial t} - \mathbf{F} \cdot \frac{\partial \mathbf{F}^*}{\partial t} \right) = \mathbf{F} \cdot \mathbf{F}^*\,, \tag{21}$$

where **j** is the energy current density and ρ is the energy density for the photon, $\nabla = \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}$. The photon propagation phase factor is $e^{-i\phi}$ in this case (see next section) and $\frac{\partial \mathbf{F}}{\partial t} = i\omega \mathbf{F}$. The energy density ρ is positive.

Now the angular momentum increment for the photon is

$$d\mathbf{S} = \hat{\mathbf{z}} \left(x j_y - y j_x \right) \frac{dV}{c^2}\,, \tag{22}$$

where **S** is the angular momentum vector or spin for the photon, j_x/c^2 and j_y/c^2 are the momentum densities in the x and y directions, respectively, and dV is the tiny volume in space. Notice that **j** needs to be divided by c^2 to be converted to the momentum density. The angular momentum for the photon in the present case has only the z component and zero x and y components. Eq. (22) may be rewritten in the form of spin momentum density as

$$\frac{dS}{dV} = \frac{1}{c^2} \left(x j_y - y j_x \right). \tag{23}$$

In the next section we present results using relationships developed here and also provide discussions on the results.

3 Results and Discussions

To start this section we first present the mathematical solution of the four-potential for the polarized photon, which are two traveling wave functions, one for the electric potential ψ, which is a scalar, and the other for the vector potential **A**. These functions are desirable since they are limited in space and show wave-particle duality with a limited length. These basic representations are important since, from which we may derive other physical quantities for the photon, such as electromagnetic fields and the spin angular momentum.

Now the solution for the four-potential in elliptic cylindrical coordinates is

$$\psi = \psi_0 \, e^{-\mu} \sin(\phi) \tag{24}$$

and

$$\mathbf{A} = \hat{\mathbf{z}} A_0\, e^{-\mu} \sin(\phi)\,, \tag{25}$$

where we assume that the photon travels in the z direction.

The vector potential in this case has only a z component. The choice of the sine function here is arbitrary, one may use the cosine function but the result should be similar since they only have a phase difference of $\pi/2$. Notice that (24) and (25) are in the same form with corresponding magnitude, and with the same phase change in both space and time. Since the physical meaning of the electric potential ψ is clear, $c^2\mathbf{A}$ may be interpreted as an electric potential current or the total-electric-potential current density flowing in the same direction as the photon, which satisfy the continuity equation given by the Lorenz gauge condition (3). Hence the Lorenz gauge may be considered as the conservation of the total-electric-potential, a physical quantity of the integration of electric potential in the whole space. With the Lorenz gauge, we can get the relationship between the two constants as in (8)

Comparing with that of circularly polarized photons [5], the strength of the four-potential for the elliptically polarized photon decreases exponentially with μ in the single space region, while the other is divided into two regions by a parameter r_0 and decreases with $1/r$ for $r > r_0$, where r is the radial value in polar cylindrical coordinates. As a result, the potential strength for the polarized photon with certain energy decreases quicker with distance from its center than that for circularly polarized photon, and hence the polarized photon may occupy less space.

Now we present expressions for the electric and magnetic fields using (9) and (10):

$$\mathbf{E} = \frac{\psi_0\, e^{-\mu}}{a\gamma}\,[\hat{\mu}\sin(\phi) - \hat{\nu}\cos(\phi)]\,, \tag{26}$$

and

$$\mathbf{B} = \frac{A_0\, e^{-\mu}}{a\gamma}\,[\hat{\mu}\cos(\phi) + \hat{\nu}\sin(\phi)]\,. \tag{27}$$

These results of $\mathbf{E}$ and $\mathbf{B}$ show that they are transverse waves and are perpendicular to each other. The energy density in classical theory for the photon is

$$\rho = \frac{1}{2}\left(\epsilon\mathbf{E}^2 + \frac{\mathbf{B}^2}{u}\right) = \frac{\epsilon{\psi_0}^2}{a^2\gamma^2}\, e^{-2\mu} \tag{28}$$

and the Poynting vector is

$$\mathbf{P} = \frac{\mathbf{E}\times\mathbf{B}}{u} = \hat{\mathbf{z}}\,\frac{c\epsilon{\psi_0}^2}{a^2\gamma^2}\, e^{-2\mu}\,, \tag{29}$$

where, in converting A_0, we used (8). These quantities are finite in space and are physically meaningful. The magnitudes of these quantities decrease exponentially with 2μ. Since the factor $a^2\gamma^2$ is equal to the combination of scale factors for both μ and ν, it can be canceled in each space integration by the same volume factor as shown later. With the Poynting vector, the photon may be viewed as a packet of energy moving at the speed of light along its propagation direction.

Since a photon is actually a quantum entity in modern physics view, we need an integral expression as (11). This is a linear combination of both the electric field and magnetic field for the elliptically polarized photon. Therefore we have a photon wave function. There are at least two advantages to have the wave function. First it can be used to calculate the value of the angular momentum for the photon; secondly it may be used to calculate the penetration probability for the photon in a sub-wavelength hole since in the view of quantum mechanics it represents the photon probability distribution. But in this article, we aim at the angular momentum value for the photon with the wave function.

In the following, we first obtain an explicit wave function using the developed expression in last section, (11), secondly derive the component expressions for energy current densities, and finally calculate the angular momentum value for the photon. This procedure has been first applied successfully to the circularly polarized photon. In this article, we report results on elliptically polarized photon.

By inserting results from (26) and (27) into (11), we have a photon wave function:

$$\mathbf{F} = \frac{\sqrt{\epsilon}\psi_0\, e^{-i\phi}}{\sqrt{2}a\gamma} e^{-\mu}(i\hat{\mu} - \hat{\nu})\,. \tag{30}$$

Using the unit vector conversion (15), we have the cartesian components of $\mathbf{F}$ as

$$F_x = \frac{\sqrt{\epsilon}\psi_0\, e^{-i\phi}}{\sqrt{2}a\gamma^2}\, e^{-\mu}(i\sinh\mu\cos\nu + \cosh\mu\sin\nu)\,, \tag{31}$$

$$F_x{}^* = \frac{\sqrt{\epsilon}\psi_0\, e^{i\phi}}{\sqrt{2}a\gamma^2}\, e^{-\mu}(-i\sinh\mu\cos\nu + \cosh\mu\sin\nu)\,, \tag{32}$$

$$F_y = iF_x\,, \tag{33}$$

and F_z is zero.

Due to the simple relationship between F_y and F_x, we have

$$F_y{}^*\,\nabla F_y = F_x{}^*\,\nabla F_x \tag{34}$$

and

$$F_y\,\nabla F_y{}^* = F_x\,\nabla F_x{}^*\,. \tag{35}$$

Hence in this case, (20) becomes

$$\mathbf{j} = \frac{c^2}{i\omega}\,(F_x\,\nabla F_x{}^* - F_x{}^*\,\nabla F_x) \tag{36}$$

and the work is reduced to one component. Furthermore since

$$F_x\,\nabla F_x{}^* = (F_x{}^*\,\nabla F_x)^*\,, \tag{37}$$

we have

$$\mathbf{j} = -\frac{2c^2}{\omega}\,\mathrm{Im}(F_x{}^*\,\nabla F_x)\,, \tag{38}$$

where “Im” means taking the real value of the imaginary part. And similarly, (21) becomes

$$\rho = \mathbf{F} \cdot \mathbf{F}^{*} = 2F_x \cdot {F_x}^{*}\,. \tag{39}$$

Now insertimg (31) and (32) into (39), we have

$$\rho = \frac{\epsilon {\psi_0}^2}{a^2\gamma^2}\, e^{-2\mu}\,, \tag{40}$$

which is the same as that of (28) for photon energy density. Now we do integration of (40) in space with the tiny volume, $dV = a^2\gamma^2 d\mu d\nu dz$. Assuming the photon length is $n\lambda$, where λ is the wavelength of the photon and n may be a positive integer, but is not exactly determined in the present work. The result should be equal to the photon energy $\hbar\omega$, where $\hbar$ is the reduced Planck constant. By doing that, we determine the electric potential constant to be

$$\psi_0 = \sqrt{\frac{2\hbar c}{\epsilon n}\,\frac{1}{\lambda}}\,. \tag{41}$$

Now we evaluate the energy current densities for the photon. F_x contains z explicitly in ϕ of the exponential function, therefore the derivative with z is simple. We have $\frac{\partial F_x}{\partial z} = -ikF_x$ and

$$j_z = \frac{c\epsilon {\psi_0}^2}{a^2\gamma^2}\, e^{-2\mu}\,, \tag{42}$$

which is consistent with the Poynting vector (29).

And from (38), we have

$$j_x = -\frac{2c^2}{\omega}\, \mathrm{Im}\left({F_x}^{*}\, \frac{\partial F_x}{\partial x}\right) \tag{43}$$

and

$$j_y = -\frac{2c^2}{\omega}\, \mathrm{Im}\left({F_x}^{*}\, \frac{\partial F_x}{\partial y}\right). \tag{44}$$

The work is now turned to calculate $\frac{\partial F_x}{\partial x}$ and $\frac{\partial F_x}{\partial y}$. Because F_x contains explicit variables of μ and ν, we need the following equations to calculate the cartesian derivatives,

$$\frac{\partial F_x}{\partial x} = \frac{\partial F_x}{\partial \mu}\,\frac{\partial \mu}{\partial x} + \frac{\partial F_x}{\partial \nu}\,\frac{\partial \nu}{\partial x} \tag{45}$$

and

$$\frac{\partial F_x}{\partial y} = \frac{\partial F_x}{\partial \mu}\,\frac{\partial \mu}{\partial y} + \frac{\partial F_x}{\partial \nu}\,\frac{\partial \nu}{\partial y}\,, \tag{46}$$

where $\frac{\partial \mu}{\partial x}, \frac{\partial \nu}{\partial x}, \frac{\partial \mu}{\partial y}, \frac{\partial \nu}{\partial y}$ may be obtained from (16). We find that

$$\begin{aligned}\frac{\partial F_x}{\partial \mu} = \beta\bigg[i\left(\cosh\mu - \sinh\mu - 2\frac{\sinh^2\mu\cosh\mu}{\gamma^2}\right)\cos\nu + \\ + \left(\sinh\mu - \cosh\mu - 2\frac{\sinh\mu\cosh^2\mu}{\gamma^2}\right)\sin\nu\bigg],\end{aligned} \tag{47}$$

$$\begin{aligned}\frac{\partial F_x}{\partial \nu} = \beta\bigg[- i\left(\cosh\mu + \sinh\mu + 2\frac{\sinh\mu\cos^2\nu}{\gamma^2}\right)\sin\nu + \\ + \left(\cosh\mu + \sinh\mu - 2\frac{\cosh\mu\sin^2\nu}{\gamma^2}\right)\cos\nu\bigg],\end{aligned} \tag{48}$$

where $\beta = \sqrt{\epsilon}\psi_0 e^{-\mu} e^{-i\phi} / \sqrt{2}a\gamma^2$.

Now the cartesian derivatives are

$$\begin{aligned}\frac{\partial F_x}{\partial x} = \beta'\bigg[i\bigg(& \cosh^2\mu\sin^2\nu - \\ & - \sinh^2\mu\cos^2\nu + \sinh\mu\cosh\mu - \\ & - 2\sinh\mu\cosh\mu\cos^2\nu\frac{\sinh^2\mu - \sin^2\nu}{\gamma^2}\bigg) - \\ & - \sin\nu\cos\nu\bigg(1 + 2\sinh\mu\cosh\mu + \\ & + 2\cosh^2\mu\,\frac{\sinh^2\mu - \sin^2\nu}{\gamma^2}\bigg)\bigg],\end{aligned} \tag{49}$$

$$\begin{aligned}\frac{\partial F_x}{\partial y} = \beta'\bigg[& i\sin\nu\cos\nu\bigg(1 - 2\sinh\mu\cosh\mu - \\ & - 2\sinh^2\mu\frac{\cosh^2\mu + \cos^2\nu}{\gamma^2}\bigg) + \\ & + \sinh^2\mu\cos^2\nu - \cosh^2\mu\sin^2\nu + \sinh\mu\cosh\mu - \\ & - 2\sinh\mu\cosh\mu\sin^2\nu\,\frac{\cosh^2\mu + \cos^2\nu}{\gamma^2}\bigg],\end{aligned} \tag{50}$$

where $\beta' = \sqrt{\epsilon}\psi_0 e^{-\mu} e^{-i\phi} / \sqrt{2}a^2\gamma^4$. These expressions are a little bit long but manageable. The purpose here is to serve as check points to guide the reader to the final correct results.

Using (43) and (44), we have

$$j_x = -\beta''\sin\nu\left(\cosh\mu + \sinh\mu\frac{\cosh^2\mu + \cos^2\nu}{\gamma^2}\right), \tag{51}$$

$$j_y = \beta''\cos\nu\left(\sinh\mu + \cosh\mu\frac{\sinh^2\mu - \sin^2\nu}{\gamma^2}\right), \tag{52}$$

where $\beta'' = c^2\epsilon{\psi_0}^2 e^{-2\mu}/\omega a^3\gamma^4$.

Now using (23), we have

$$\begin{aligned}\frac{dS}{dV} = \alpha e^{-2\mu}\bigg(& \frac{\sinh\mu\cosh\mu}{\gamma^2} + \\ & + \frac{\sinh^2\mu\cosh^2\mu - \sin^2\nu\cos^2\nu}{\gamma^4}\bigg),\end{aligned} \tag{53}$$

where $\alpha = \epsilon{\psi_0}^2/\omega a^2\gamma^2$.

To calculate the spin value, we integrate (53) in the whole space. There are two parts to be integrated on the right hand side of the equation. This integration is a bit challenging since each integration part is divergent at $\mu = 0$ and $\nu = 0, \pi$. To avoid this problem we work around by first doing the integration of the second part which fortunately produces an exact term to cancel the first part and the remaining is finite and manageable. We now show the integration of the second part:

$$\begin{aligned} I = &\int_0^{n\lambda}\int_0^{\infty}\int_0^{2\pi} dz\, d\mu\, d\nu \\ &e^{-2\mu}\,\frac{\sinh^2\mu\cosh^2\mu-\sin^2\nu\cos^2\nu}{\gamma^4} \\ = n\lambda &\int_0^{\infty}\int_0^{2\pi} d\mu\, d\nu \\ &e^{-2\mu}\,\frac{\sinh^2\mu\cosh^2\mu-\sin^2\nu\cos^2\nu}{\gamma^4}\,, \end{aligned} \tag{54}$$

where the scale factors in the integration volume are canceled within the α factor and we omit the rest of the constants here for simplicity. This integration may be further separated into sub-integration as

$$I_1 = \int_0^{\infty}\int_0^{2\pi} e^{-2\mu}\,\frac{\sinh^2\mu\cosh^2\mu}{(\sinh^2\mu+\sin^2\nu)^2}\,d\mu\, d\nu \tag{55}$$

and

$$I_2 = \int_0^{\infty}\int_0^{2\pi} e^{-2\mu}\,\frac{\sin^2\nu\cos^2\nu}{(\sinh^2\mu+\sin^2\nu)^2}\,d\mu\, d\nu\,. \tag{56}$$

These may be done by the partial integration method: for (55) first integrate with μ and for (56) first integrate with ν. Hence we have

$$\begin{aligned} I_1 = &-\int_0^{\infty}\int_0^{2\pi} e^{-2\mu}\,\frac{\sinh\mu\cosh\mu}{\sinh^2\mu+\sin^2\nu}\,d\mu\, d\nu + \\ &+\frac{1}{2}\int_0^{\infty}\int_0^{2\pi} e^{-2\mu}\,\frac{\sinh^2\mu+\cosh^2\mu}{\sinh^2\mu+\sin^2\nu}\,d\mu\, d\nu - \\ &-\frac{1}{2}\left[e^{-2\mu}\,\sinh\mu\cosh\mu\int_0^{2\pi}\frac{d\nu}{\sinh^2\mu+\sin^2\nu}\right]\Bigg|_0^{\infty} \end{aligned} \tag{57}$$

and

$$I_2 = \frac{1}{2}\int_0^{\infty}\int_0^{2\pi} e^{-2\mu}\,\frac{\cos^2\nu-\sin^2\nu}{\sinh^2\mu+\sin^2\nu}\,d\mu\, d\nu\,. \tag{58}$$

Now the last integration term in (57) is zero at both $\mu = 0$ and $\mu \to \infty$. Hence (54) becomes

$$\begin{aligned} I = &-n\lambda\int_0^{\infty}\int_0^{2\pi} e^{-2\mu}\,\frac{\sinh\mu\cosh\mu}{\sinh^2\mu+\sin^2\nu}\,d\mu\, d\nu + \\ &+n\lambda\int_0^{\infty}\int_0^{2\pi} e^{-2\mu}\,d\mu\, d\nu\,, \end{aligned} \tag{59}$$

where the second integration term is the second integration term of (57) minus that of (58). And finally by finishing the second integration we have

$$I = -\int_0^{n\lambda}\int_0^{\infty}\int_0^{2\pi} e^{-2\mu}\,\frac{\sinh\mu\cosh\mu}{\gamma^2}\,dz\, d\mu\, d\nu + n\lambda\pi. \tag{60}$$

The first integration term in (60) cancels exactly the integration of the first part in (53) so the angular momentum for the photon is

$$S = \frac{\epsilon{\psi_0}^2}{\omega}\,n\lambda\pi = \hbar\,, \tag{61}$$

where we used (41). The value of spin or the angular momentum calculated here for the elliptically polarized photon is indeed $\hbar$.

Before concluding this section we consider the elliptic coordinate parameter a for the photon. The divergence of the electric field (26) is zero everywhere except at the two focal points ($x = \pm a$). This leads us to believe that electricity may only exist in these two focal points formed traveling lines. To further consider the value of a we take a look at that for circularly polarized photon [5]. In that case the electromagnetic field occupies two space regions divided by r_0 with the center core region carrying zero angular momentum for spin one. The elliptically polarized photon may be understood as transformed from the circularly polarized photon with its core region collapsed by its energy popped out without change in its length of circumference. If that is the case then $a = \lambda/4$.

4 Conclusion

To conclude this article we summarize what has been presented here. First, we have solved the wave equations for the electromagnetic four-potential in the elliptic cylindrical coordinates for the polarized photon. The solution for each potential is an electromagnetic traveling wave and its transverse strength decreases exponentially with μ. These expressions for the four-potential are simple but essential representations since they may be used to obtain other physical quantities for the polarized photon.

We first obtained the electric field and magnetic field for the photon from the four-potential solution. Then we have presented the energy wave function explicitly, which is a linear combination of the electric field and magnetic field. Using concepts from quantum mechanics, we first derived expressions then evaluated for photon energy, energy current, and angular momentum densities. Work is shown particularly in calculating the value of the angular momentum or spin for the photon. Considerations are given about the value of the elliptic coordinate parameter a which may be equal to a quarter of the photon wavelength.

Received on October 28, 2020

References

1. Jackson J. D. Classical Electrodynamics, 3rd ed. John Wiley & Sons Inc., New York, 1999.
2. Einstein A. *Annalen der Physik* (in German), 1905, v. 17 (6), 132. Einstein A. On a Heuristic Point of View Concerning the Production and Translation of Light. The Swiss Years: Writings, 1900–1909, v. 2, 86.
3. Millette P. A. Wave-Particle Duality in the Elastodynamics of the Spacetime Continuum (STCED). *Progress in Physics*, 2014, v. 10 (4), 255–258.

4. Lehnert B. Joint Wave-Particle Properties of the Individual Photon. *Progress in Physics*, 2007, v. 4 , 104–108.
5. Weng S. A Classical Model of the Photon. *Progress in Physics*, 2016, v. 12 (1), 49–55.
6. Gordon R., Brolo A. G., McKinnon A., Rajora A., Leathem B., and Kavanagh B. L. Strong Polarization in the Optical Transmission through Elliptical Nanohole Arrays. *Physical Review Letters*, 2004, v. 92 (3), 037401(4).
7. Arfken G., Mathematical Methods for Physicists, 3rd ed. Academic Press, Inc., New York, 1985
8. Bialynicki-Birula I. Photon Wave Function. In Wolf E., ed. Progress in Optics XXXVI. Elsevier, Amsterdam, 1996, 245–294. arXiv: quantph/0508202v1.
9. Davydov A. S. Quantum Mechanics, 2nd ed. Translated by Ter Haar D., Pergamon Press, New York, 1988

A Note on the Barut Second-Order Equation

Valeriy V. Dvoeglazov

UAF, Universidad Autónoma de Zacatecas, Apartado Postal 636, Zacatecas 98061 Zac., México
E-mail: valeri@fisica.uaz.edu.mx

The second-order equation in the $(1/2, 0) \oplus (0, 1/2)$ representation of the Lorentz group has been proposed by A. Barut in the 70s [1]. It permits to explain the mass splitting of leptons (e, μ, τ). The interest is growing in this model (see, for instance, the papers by S. Kruglov [2] and J. P. Vigier *et al.* [3, 4]). We note some additional points of this model.

The Barut main equation is

$$\left[i\gamma^\mu \partial_\mu + \alpha_2 \partial^\mu \partial_\mu - \kappa\right] \Psi = 0\,, \tag{1}$$

where α_2 and κ are the constants later related to the anomalous magnetic moment and mass, respectively. The matrices γ^μ are defined by the anticommutation relation:

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}\,, \tag{2}$$

$g^{\mu\nu}$ is the metrics of the Minkowski space, $\mu, \nu = 0, 1, 2, 3$. The equation represents a theory with the conserved current that is linear in 15 generators of the 4-dimensional representation of the O(4, 2) group, $N_{ab} = \frac{i}{2}\gamma_a \gamma_b, \gamma_a = \{\gamma_\mu, \gamma_5, i\}$. Instead of 4 solutions, (1) has 8 solutions with the correct relativistic relation $E = \pm\sqrt{\mathbf{p}^2 + m_i^2}$. In fact, it describes states of different masses (the second one is $m_2 = 1/\alpha_2 - m_1 = m_e(1 + 3/2\alpha)$, α is the fine structure constant), provided that the certain physical condition is imposed on $\alpha_2 = (1/m_1)(2\alpha/3)/(1 + 4\alpha/3)$, the parameter (the anomalous magentic moment should be equal to $4\alpha/3$). One can also generalize the formalism to include the third state, the τ-lepton [1b]. Barut has indicated the possibility of including γ_5 terms (e.g. $\sim \gamma_5 \kappa'$).

The most general form of spinor relations in the $(1/2, 0) \oplus (0, 1/2)$ representation has been given by Dvoeglazov [5]. It was possible to derive the Barut equation from first principles [6]. Let us reveal the connections with other models. For instance, in [3, 7] the following equation has been studied:

$$\begin{gathered}\left[\left(i\hat{\partial} - e\hat{A}\right)\left(i\hat{\partial} - e\hat{A}\right) - m^2\right] \Psi = \\ \left[\left(i\partial_\mu - eA_\mu\right)(i\partial^\mu - eA^\mu) - \tfrac{1}{2} e\sigma^{\mu\nu}F_{\mu\nu} - m^2\right] \Psi = 0\end{gathered} \tag{3}$$

for the 4-component spinor Ψ. $\hat{A} = \gamma^\mu A_\mu$; A_μ is the 4-vector potential; e is electric charge; $F_{\mu\nu}$ is the electromagnetic tensor. $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]_-$. This is the Feynman-Gell-Mann equation. In the free case we have the Lagrangian (see Eq. (9) of [3c]):

$$\mathcal{L}_0 = (\overline{i\hat{\partial}\Psi})(i\hat{\partial}\Psi) - m^2\overline{\Psi}\Psi\,. \tag{4}$$

Let us re-write (1) into the form:*

$$\left[i\gamma^\mu \partial_\mu + a\partial^\mu \partial_\mu + b\right] \Psi = 0\,. \tag{5}$$

So, one should calculate ($p^2 = p_0^2 - \mathbf{p}^2$)

$$\mathrm{Det}\begin{pmatrix} b - ap^2 & p_0 + \boldsymbol{\sigma}\cdot\mathbf{p} \\ p_0 - \boldsymbol{\sigma}\cdot\mathbf{p} & b - ap^2 \end{pmatrix} = 0 \tag{6}$$

in order to find energy-momentum-mass relations. Thus, $[(b - ap^2)^2 - p^2]^2 = 0$ and if $a = 0$, $b = \pm m$ we come to the well-known relation $p^2 = p_0^2 - \mathbf{p}^2 = m^2$ with four Dirac solutions. However, in the general case $a \neq 0$ we have

$$p^2 = \frac{(2ab + 1) \pm \sqrt{4ab + 1}}{2a^2} > 0\,, \tag{7}$$

that signifies that we do not have tachyons. However, the above result implies that we cannot just put $a = 0$ in the solutions, while it was formally possible in (5). When $a \to 0$ then† $p^2 \to \infty$; when $a \to \pm\infty$ then $p^2 \to 0$. It should be stressed that *the limit in the equation does not always coincide with the limit in the solutions.* So, the questions arise when we consider limits, such as Dirac → Weyl, and Proca → Maxwell. The similar method has also been presented by S. Kruglov for bosons [8]. Other fact should be mentioned: when $4ab = -1$ we have only the solutions with $p^2 = 4b^2$. For instance, $b = m/2$, $a = -1/2m$, $p^2 = m^2$. Next, I just want to mention one Barut omission. While we can write

$$\frac{\sqrt{4ab + 1}}{a^2} = m_2^2 - m_1^2\,, \text{ and } \quad \frac{2ab + 1}{a^2} = m_2^2 + m_1^2\,, \tag{8}$$

but m_2 and m_1 should not necessarily be associated with $m_{\mu,e}$ (or $m_{\tau,\mu}$). They may be associated with their superpositions, and applied to neutrino mixing, or quark mixing.

The lepton mass splitting has also been studied by Markov [9] on using the concept of both positive and negative masses in the Dirac equation. Next, obviously we can calculate anomalous magnetic moments in this scheme (on using, for instance, methods of [10, 11]).

We previously noted:

- The Barut equation is a sum of the Dirac equation and the Feynman-Gell-Mann equation.
- Recently, it was suggested to associate an analogue of (4) with dark matter, provided that Ψ is composed of

---

*Of course, one could admit p^4, p^6 etc. in the Dirac equation too. The dispersion relations will be more complicated [6].

†a has dimensionality [1/m], b has dimensionality [m].

the self/anti-self charge conjugate spinors, and it has the dimension $[\text{energy}]^1$ in the unit system $c = \hbar = 1$. The interaction Lagrangian is $\mathcal{L}^H \sim g\bar{\Psi}\Psi\phi^2$, ϕ is a scalar field.

- The term $\sim \overline{\Psi}\sigma^{\mu\nu}\Psi F_{\mu\nu}$ will affect the photon propagation, and non-local terms will appear in higher orders.
- However, it was shown in [3b,c] that a) the Mott cross-section formula (which represents the Coulomb scattering up to the order $\sim e^2$) is still valid; b) the hydrogen spectrum is not much disturbed; if the electromagnetic field is weak the corrections are small.
- The solutions are the eigenstates of the γ^5 operator.
- In general, the current J_0 is not the positive-defined quantity, since the general solution $\Psi = c_1\Psi_+ + c_2\Psi_-$, where $[i\gamma^\mu\partial_\mu \pm m]\Psi_\pm = 0$, see also [9].
- We obtained the Barut-like equations of the 2nd order and 3rd order in derivatives.
- We obtained dynamical invariants for the free Barut field on the classical and quantum level.
- We found relations with other models (such as the Feynman-Gell-Mann equation).
- As a result of analysis of dynamical invariants, we can state that at the free level, the term $\sim \partial_\mu\overline{\Psi}\sigma_{\mu\nu}\partial_\nu\Psi$ in the Lagrangian does not contribute.
- However, the interaction terms $\sim \bar{\Psi}\sigma_{\mu\nu}\partial_\nu\Psi A_\mu$ will contribute when we construct the Feynman diagrams and the S-matrix. In the curved space (the 4-momentum Lobachevsky space), the influence of such terms has been investigated in the Skachkov work [10, 11]. Briefly, the contribution will be such as if the 4-potential were to interact with some "renormalized" spin. Perhaps, this explains why Barut used the classical anomalous magnetic moment $g \sim 4\alpha/3$ instead of $\alpha/2\pi$.

Acknowledgements

The author acknowledges discussions with participants of recent conferences.

Received on November 17, 2020

References

1. Barut A. O. *Phys. Lett.*, 1978, v. B73, 310; *Phys. Rev. Lett.*, 1979, v. 42, 1251; Wilson R. *Nucl. Phys.*, 1974, v. B68, 157.
2. Kruglov S. I. *Ann. Fond. Broglie*, 2004, v. 29 (H2), 1005. In: Dvoeglazov V. V. et al., eds. The special issue dedicated to Yang and Mills. arXiv: quant-ph/0408056.
3. Petroni N. C., Vigier J. P. et al. *Nuovo Cim.*, 1984, v. B81, 243; *Phys. Rev.*, 1984, v. D30, 495; ibid, 1985, v. D31, 3157.
4. Dvoeglazov V. V. *J. Phys. Conf. Ser.*, 2005, v. 24, 236–240; *Adv. Appl. Clifford Algebras*, 2008, v. 18, 579–585.
5. Dvoeglazov V. V. *Hadronic J. Suppl.*, 1995, v. 10, 349; *Int. J. Theor. Phys.*, 1998, v. 37, 1909.
6. Dvoeglazov V. V. *Ann. Fond. Broglie*, 2000, v. 25, 81–92.
7. Feynman R. and Gell-Mann M. *Phys. Rev.*, 1958, v. 109, 193.
8. Kruglov S. I. *Int. J. Mod. Phys.*, 2001, v. A16, 4925–4938. arXiv: hep-th/0110083.
9. Markov M. *ZhETF*, 1937, v. 7, 579; ibid, 1937, v. 7, 603; *Nucl. Phys.*, 1964, v. 55, 130.
10. Skachkov N. B. *Theor. Math. Phys.*, 1975, v. 22, 149; ibid, 1976, v. 25, 1154.
11. Dvoeglazov V. V. and Skachkov N. B. *Sov. J. Nucl. Phys.*, 1988, v. 48, 1065.

The Unification of the Basic Units Meter, Kilogram and Second and the Essence of the Phenomenon Time

Gerhard Dorda

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

It is shown by experimental data that a causal connection between the categories mass and time, as well as between the categories electric current and time is given. The equation of the mass–time relation, valuated together with the constant of the velocity of light c and the length–mass relation of Planck L_E/M_E, results in a specific, single number related equation of the units mass, length and time without any dimension, thus representing unreality. It is made evident that the unified equation of the basic units, which reflects the not explicable experimental findings of the Quantum-Hall-Effect (QHE, i.e. KE), the findings of the physical description of vision and sound and also the third law of Kepler, yields the possibility to describe the essence of time. It is shown that the Hubble time $T_{U,E}$ and the Earth-related time t_E should be considered to be the fundamental factor of realization of masses between unreality and reality. Based on the presented description of the essence of the phenomenon time, the difference between time and frequency is disclosed.

1 Introduction

The MOS transistors as an amplifier of electric signals was developed after the second world war in the USA by W. Shockley, W. H. Brattain and J. Bardeen. Its economic importance is given by its extraordinary ability of miniaturization. This fact was the start of a world-wide rapid technical development at all areas of economy. The extensive studies of the specific properties of these MOS transistors led to the observation of the Quantum-Hall-Effect (QHE) in 1980 by K. von Klitzing, thus named Klitzing-Effect (KE) [1]. This effect, observed at low temperatures, disclosed on the one hand the existence of a macroscopic quantization at discrete states, given by the quantization of resistivity of the MOS module in form of $R_{xx} = h/ie^2$ (here, h is the Planck-constant, e the charge of electron, and i the quantum-number), and on the other hand the existence of a simultaneous, i.e. contemporaneous condition of an unresistance $R_{xx} = 0$ ohm. Both these effects, especially the contemporaneity of $R_{xx} = h/ie^2$ and $R_{xx} = 0$ ohm, was at that time not foreseen by the given theory, i.e. also quantum mechanical theory.

The exceptional importance of the QHE is given not only by the observation of a macroscopical quantization, described by the $R_{xx} = h/ie^2$, but also by the unexpected finding of the simultaneously given independence $R_{xx} = 0$ ohm at any integral and fractional quantum number. Evidently, the observed $R_{xx} = h/ie^2$ effect yields the possibility to find a connection to the state of physics before the year 1980, the year of the observation of the QHE by K. von Klitzing (KE) [1], but the $R_{xx} = 0$ ohm effect is a quite new observation within the whole scientific field of being, showing the existence of a state of space–mass–time independence. This finding has been observed after a world-wide extensive experimental investigation of the QHE. This spectacular observation allows to postulate that due to the independence of the $R_{xx} = 0$ ohm effect on the integral as well as fractional quantum number that especially this $R_{xx} = 0$ ohm effect represents the fundamental background of the QHE. This assumption will be confirmed in the following by the description of the essence of time, especially by the unforeseen formulation of an equation of space–mass–time independence.

The surprising observation of the existence of a state without length, mass and time suggests a reform of the *International System of Units* (SI). It should be pointed out that a reform of the SI was recently highly recommended by F. W. Hehl and C. Lämmerzahl [2] with reference to physical results observed in the last centuries, and thus also to the experimentally observed dependence of the value of the velocity of light c on gravity. In [2] it was not borne in mind that the value of c is given by a free choice [3], i.e. the numerical value of c was determined by man, and "not by nature". Thus the free choice of the numerical value of c determines in the last consequence the numerical value of the fundamental physical constants. Moreover, when describing the effect of gravitation on the number of velocity of light c we have to assume that in agreement with the physical interpretation of vision and sound [4, 5] the light-related distance refers to the localized wavy 2D-state, thus including no gravity effects, whereas the length-, i.e. mass-related distance, according to the third law of Kepler, i.e. due to the three-dimensionality, includes also the gravitational effects: Thus the different background of the light-related distance and the gravity, i.e. mass-related distance shows that the variability of the number of the velocity of light c, caused by gravity, cannot influence the numerical values of basic units. Furthermore, as will be demonstrated in sections 4 and 5, these discoveries result in the general validity of the following discussed equations and shows that this finding reveals the applicability of all explored fundamental constants at any place of the cosmos.

Unfortunately, also the weighty problem of the phenomenon time was not incorporated in the extensive analysis of [2], though Lee Smolin [6] has shown in *The Trouble with Physics* that the main open question of the existing physics refers to not knowing of the essence of the phenomenon time. Therefore, in the presented analysis of the curious experimental effect of an existence of a state, being independent of a mass, space and time, as well as of the lack of a basic interpretation of the phenomenon time, it will be demonstrated that both these problems of physics can be solved only together. It will be shown in the next sections that the description of these phenomena does not become possible before the demonstration of the unexpectedly given causal connection between the category mass and frequency, as well as also between the category electric current and frequency.

2 The background for the basic interpretation of time

To find a physical answer to the fundamental problems of physics, we start with the analysis of the phenomenon time. It will be shown that the physical description of the process of vision and sound [4, 5], which demonstrates the existence of a differentiation of the three-dimensional space DSS into the two-dimensionality 2D, i.e. electromagnetism, and the one-dimensionality 1D, i.e. length, i.e. gravitation, is extremely helpful for the analysis of the phenomenon time. It has become aware that the description of time is not dependent on the wavelength of the used light, but solely on the frequency, reflecting the electromagnetism. Based on the DSS-model, in fact the source of the phenomenon time must refer to dynamics, i.e. to the electromagnetism, which is an effect of the 2D-state. The unit time is always noticeable merely in connection with the category length [6]. This finding suggests that the perception of the phenomenon time must be a result of the connection of the wave related 2D-state with the real, i.e. 1D-state, which represents the state of observable facts. The analysis of the process of vision has shown [4] that the 1D-state, representing the gravity and thus also the category length, relates to the effect of the gravitational constant $G_{\mathrm{E}} = c^2(L_{\mathrm{E}}/M_{\mathrm{E}})$, where L_{E} is the Planck-length and M_{E} the Planck-mass and c the velocity of light. It is generally assumed that the gravitational constant G_{E} is valid at the whole cosmos. From this decisive supposition follows that the relation $L_{\mathrm{E}}/M_{\mathrm{E}}$, i.e. the fundamental connection between the three-dimensionally related mass and the category length, is describable in an extended form, given by

$$\frac{L_{\mathrm{E}}}{M_{\mathrm{E}}} = \frac{\lambda_{\mathrm{C}}}{M_{0,\mathrm{E}}} = \frac{\lambda_{\mathrm{G,E}}}{M_{\mathrm{G,E}}} = \frac{L_{\mathrm{U,E}}}{M_{\mathrm{U,E}}} \,. \tag{1}$$

Here λ_{C} is the Compton wave-length and $M_{0,\mathrm{E}}$ the corresponding to this length related mass, $\lambda_{\mathrm{G,E}}$ is the so-called reference length of the earth and $M_{\mathrm{G,E}}$ the corresponding mass of the earth, and finally $L_{\mathrm{U,E}}$ the length of the cosmos and $M_{\mathrm{U,E}}$ its mass. The index E indicates that the explorations are performed from an earth-related place.

It should be emphasized that (1) and thus also the relation $\lambda_{\mathrm{C}}/M_{0,\mathrm{E}}$ represent the particular state of identity of the observable electromagnetism with gravitation. Considering the model of the differentiated structure of space DSS [4, 5], this state is given by the non-possibility to distinguish between the 2D-state and the 1D-state. But this specific condition does not exist at the surface of the earth. The third law of Kepler [4], which is related to the spatial three-dimensionality of the cosmos, shows that the relation between the square velocity of light c^2 and the square of the orbital speed at the surface of the earth $v_{\mathrm{E}}{}^2$ (which reflects the difference between the electromagnetism and gravitation) is given by $c^2/v_{\mathrm{E}}{}^2 = a_{\mathrm{G,E}}$. This number $a_{\mathrm{G,E}}$, related to the surface of the earth, in [4] described by (8) and (9) and formulated in agreement with (1), is given by

$$a_{\mathrm{G,E}} = \frac{M_{\mathrm{E}}\, R_{\mathrm{G,E}}}{L_{\mathrm{E}}\, M_{\mathrm{G,E}}} = \frac{M_{0,\mathrm{E}} R_{\mathrm{G,E}}}{\lambda_{\mathrm{C}} M_{\mathrm{G,E}}} = \frac{R_{\mathrm{G,E}}}{\lambda_{\mathrm{G,E}}} \,, \tag{2}$$

where $R_{\mathrm{G,E}}$ is the radius of the earth and $M_{\mathrm{G,E}}$ its mass.

It should be pointed out that according to the DSS-model we have to proceed from the 1D, i.e. from the noticeable, i.e real state. In the next sections, it is demonstrated that, according to the spatial three-dimensionality, the 2D-state is valued in a "square" relation to the 1D-state. Thus the frequency f_{C}, being related to the 2D-state, must be modified on the surface of earth by $(a_{\mathrm{G,E}})^{1/2}$, resulting in a real value, given by $f_{\mathrm{E}} = f_{\mathrm{C}}(a_{\mathrm{G,E}})^{-1/2}$. When we take the value $f_{\mathrm{C}} = 1.235589964 \times 10^{20}$ Hz for the velocity of light related frequency, the value $M_{\mathrm{G,E}} = 5.974 \times 10^{24}$ kg for the mass of the earth and the approximate value $R_{\mathrm{G,E}} = 6.36 \times 10^6$ m [7] for its radius $R_{\mathrm{G,E}}$, then, according to (2), we obtain for f_{E} the real value

$$f_{\mathrm{E}} = f_{\mathrm{C}}\,(a_{\mathrm{G,E}})^{-1/2} = 3.26321(64) \times 10^{15}\ \mathrm{Hz}\,. \tag{3a}$$

It is of great importance for our further analysis to compare this experimentally established numerical value of f_{E} with the numerical value $M_{0,\mathrm{E}} = \lambda_{\mathrm{C}}(M_{\mathrm{E}}/L_{\mathrm{E}})$, which refers to (1). As known, the Compton-wavelength is given by

$$\lambda_{\mathrm{C}} = 2.4263102389 \times 10^{-12}\ \mathrm{m}\,,$$

and for the relation $L_{\mathrm{E}}/M_{\mathrm{E}}$ we have the value

$$L_{\mathrm{E}}/M_{\mathrm{E}} = 7.42565(74) \times 10^{-28}\ \mathrm{m\,kg^{-1}}.$$

This value of $L_{\mathrm{E}}/M_{\mathrm{E}}$ was determined from the analysis of the cosmos generally used, experimentally observed gravitation constant $G_{\mathrm{E}} = c^2(L_{\mathrm{E}}/M_{\mathrm{E}}) = 6.67384(80) \times 10^{-11}\ \mathrm{m^3/(kg\,s^2)}$ [7]. Based on (1) and using these values, we obtain for $M_{0,\mathrm{E}}$ a 1D related value, given by

$$M_{0,\mathrm{E}} = \lambda_{\mathrm{C}}\,(M_{\mathrm{E}}/L_{\mathrm{E}}) = 3.26746(86) \times 10^{15}\ \mathrm{kg}\,. \tag{3b}$$

The comparison of the "numerical" value f_{E}, given by (3a), with the "numerical" value $M_{0,\mathrm{E}}$, given by (3b), "astonishingly" reveals a near identity of their numbers. This exceptional finding leads us to the dared assumption that a causal connection between mass and frequency, i.e. time, seems to be possible in being. This exceptional assumption can be formulated by means of the spectacular equation

$$M_{0,\mathrm{E}} = f_{\mathrm{E}} \,. \tag{4}$$

The comparison of the values of $M_{0,\mathrm{E}}$ with f_{E} shows that the experimental value of f_{E} is a little higher, but only about 0.13 %. This relatively small deviation is necessarily a consequence of the fact that according to vision and sound the effect of $M_{\mathrm{G,E}}$ on the value of $a_{\mathrm{G,E}}$ is a little lower, caused by the reduced earth density at the surface. Therefore, indeed, it is physically allowed to postulate that on the surface of the earth an absolute numerical identity of $M_{0,\mathrm{E}}$ with f_{E} is given, as proposed by (4).

It should be pointed out that the careful analysis of (4), given in sections 3–5, shows that the proposed identity of the limits of mass with frequency has to be valid not only on the surface of the earth, but generally valid at the whole cosmos. Thus, it should be considered that beside the physical constants c and $L_{\mathrm{E}}/M_{\mathrm{E}}$, a third important, generally valid constant should be effective, representing a connection of mass with frequency. Due to the importance of this exceptional postulate, experimental findings will be presented in the next section to substantiate the validity of the extraordinary (4).

3 Experimental verifications of the identity of the mass–time connection

3.1 The analysis of the limiting current of the Quantum-Hall-Effect (QHE)

The limiting current of the QHE, obtained by the experimental investigation of W. Wittmann [8], is presented in Fig. 2.1 of [9], page 37. Assessing these data with respect to the process of seeing and hearing, we have to conclude that the investigated electric current of the sample, the so-called source-drain current I_{SD}, being a real effect, must be related to the 1D-state, i.e. to frequency being in the real form of time, and not to the 2D wave state. Therefore it must be concluded that the factor of modification $(a_{\mathrm{G,E}})^{1/2}$ has to be in relation solely to the source-drain frequency f_{SD} and not also to the charge of the electron e of the electric current, as former assumed in [9]. This shows, formulated in a general form, that the limiting frequency on the surface of the earth f_{E} must be given by (3a) and thus the limiting current $I_{0,\mathrm{E}}$ by

$$I_{0,\mathrm{E}} = e f_{\mathrm{E}} = 5.23510(29) \times 10^{-4}\ \mathrm{A}\,. \tag{5}$$

This theoretical value of the limiting current really agrees on the whole with the experimental data of the QHE, as shown and discussed on pages 39–40 of [9], together with the results of the extension of the time-analysis, given in [10], part II, pages 37–50. This particular finding demonstrates that the basic unit ampere of the MKSA- or SI-system of basic units must be considered to be a fix relation to the basic unit time, and that by means of the electron charge e. From this follows that the limit voltage of the QHE $V_{0,\mathrm{QHE}}$ is at the limit resistivity $R_{0,\mathrm{QHE}} = 2.581281 \times 10^4$ ohm given by $V_{0,\mathrm{QHE}} = 13.51316(38)\,\mathrm{V}$ and that the relation between the mass and the charge of the electron results in $m_{\mathrm{e}}/e = (a_{\mathrm{G,E}})^{1/2} V_{0,\mathrm{QHE}}/c^2$, which agrees with the experimental experience. These experimental results suggest the striking conclusions that, on the one side, the equation of the frequency indeed should be given by $f_{\mathrm{E}} = f_{\mathrm{C}}(a_{\mathrm{G,E}})^{-1/2}$, as proposed in section 2 and thus supporting the assumption of (4), and, on the other side, only a reduced MKS basic system of units, i.e. without the category "electric current", should be taken into account in the physical science, being a far-reaching conclusion of the QHE.

3.2 The comparison of the mass–time relation effect with the Hubble time $T_{\mathrm{U,E}}$

An indirect experimental confirmation of (4) can be obtained when we interpret the connection between the mass $M_{0,\mathrm{E}}$ and the frequency f_{E} as a fundamental coupling number, and that seen in similarity to the speed of light c, and when we consider the relation $L_{\mathrm{E}}/M_{\mathrm{E}}$ also as a fundamental coupling number. The numerical value of the relation $L_{\mathrm{E}}/M_{\mathrm{E}}$ is given by the constant of gravity G_{E} [7], where the factor c^2 is based on the value $c = 2.99792458 \times 10^8\ \mathrm{m\,s^{-1}}$ [9]. The fundamentality of the numerical value of $L_{\mathrm{E}}/M_{\mathrm{E}}$ was demonstrated by the equation of the Hubble-effect [4]. Therefore, describing the velocity of light as a fundamental coupling number between the categories length and time, given by the number 2.99792458×10^8, and the relation $L_{\mathrm{E}}/M_{\mathrm{E}}$ as a fundamental coupling number between categories length and mass, given by the number $7.42565(74) \times 10^{-28}$, thus the connection between the mass $M_{0,\mathrm{E}}$ and the frequency f_{E}, i.e. time, given by (3a), (3b) and (4), has according to our analysis to be assessed as a general valid fundamental coupling number between the relation of the categories mass and time, describable by the number 1.

Summarizing these propositions, a general valid fundamental connection between the categories length, time and mass can be achieved by

$$\begin{aligned} 1\ \mathrm{m} &= 1/c = (1/2.99792458 \times 10^8)\ \mathrm{s} \\ &= 3.335640952 \times 10^{-9}\ \mathrm{s}\,, \end{aligned} \tag{6a}$$

$$\begin{aligned} 1\ \mathrm{m} &= 1/(L_{\mathrm{E}}/M_{\mathrm{E}}) \\ &= (1/7.42565(74) \times 10^{-28})\ \mathrm{kg} \\ &= 1.346682(11) \times 10^{27}\ \mathrm{kg} \end{aligned} \tag{6b}$$

$$\text{and}\quad 1\ \mathrm{kg} = 1/(1\ \mathrm{s})^{*}\,, \tag{6c}$$

where we introduce a fundamental time $(1\,\text{s})^*$ in (6c), a consequence of the assumption of the existence of a general valid fundamental connection between the categories length, mass and time. Thus (6a)–(6c) yield

$$(1\,\text{s}^2)^* = (M_\text{E}/L_\text{E})\,c = 4.03725(14)\times 10^{35}\,\text{s}^2\,, \qquad (7)$$

being a square of the fundamental time. Thus the unit of time, representing the fundamental coupling number, therefore must be given by

$$(1\,\text{s})^* = 6.35393(69)\times 10^{17}\,\text{s}\,, \qquad (8)$$

which must be considered to be a result, related to the freely chosen value of c.

To get a further possibility to confirm the correctness of our analysis and thus of (4), we start from the idea that this specific time of (8), deduced from the numerical values c, L_E/M_E and (4), correspond with the Hubble time $T_\text{U,E}$. Throughout the scientific literature, the Hubble time $T_\text{U,E}$ is determined by means of the Hubble constant H_0, determined by telescopes. The value of the time of (8) corresponds with the Hubble-value $H_0 = 48.564\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$. Thus it is interesting that the experimentally detected H_0 values show the Hubble constant, found in the last decades, to be between $H_0 = 40\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$ and $H_0 = 100\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$ [11], and the recently determined value shows to be [4]

$$H_0 = 72.1\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}.$$

It is now clear that the values of the experimental findings of H_0, according to the size, are identical with the size of the theoretical value given by (8). Thus we can state that the size-related agreement of the telescopes given $T_\text{U,E}$ values with the theoretical value, given by (8), additionally proves that the postulated identity of the category mass with the category time, expressed by (4), indeed can be considered to be experimentally verified.

4 The formulation of an equation of transformed basic units without dimensions

As has been manifested in subsection 3.2, it is very interesting that the generally valid limit values c and L_E/M_E yield in (7) the square of the category time. As will be discussed in detail in section 5, this odd finding can be solved when we take into consideration both the experimental data of the QHE [1] and the physical description of vision and sound [4, 5].

The KE shows the $R_{xx} = 0$ ohm effect, which manifests the existence of an extraordinary state without any difference between mass, length and time. To reflect this mysterious experimental finding, a transformation of (6a)–(6c) is necessary to achieve the basic units given simply by numbers. This spectacular goal is attained by

$$1\,\text{m}^* = 2.11944(52)\times 10^{9}\,, \qquad (9a)$$

$$1\,\text{kg}^* = 1.573827(44)\times 10^{-18}\,, \qquad (9b)$$

$$1\,\text{s}^* = 6.35393(69)\times 10^{17}\,. \qquad (9c)$$

The numbers of (9a)–(9c) are obtained, when we use $1/c$ and $1/(L_\text{E}/M_\text{E})$ as fundamental coupling numbers and when we suppose that the time of (8) is identical with the limit time of the cosmos $T_\text{U,E}$. Thus the numbers of (9a)–(9c) are given by

$$1\,\text{m}^* = \frac{1}{c}\,(1\,\text{s})^*\,, \qquad (10a)$$

$$1\,\text{kg}^* = \frac{L}{M\,c}\,(1\,\text{s})^*\,, \qquad (10b)$$

$$1\,\text{kg}^* = \frac{1}{(1\,\text{s})^*}\,. \qquad (10c)$$

When starting from the cosmic length, given by $L_\text{U,E} = c\,T_\text{U,E} = 1.90486(24)\times 10^{26}$ m and from the cosmic mass, given according to (1) by

$$\begin{aligned} M_\text{U,E} &= (M_\text{E}/L_\text{E})\,L_\text{U,E} = (M_\text{E}/L_\text{E})\,c\,T_\text{U,E} \\ &= 2.56524(41)\times 10^{53}\,\text{kg}\,, \end{aligned}$$

and multiplying these values with the transformed basic units $1\,\text{m}^*$ and $1\,\text{kg}^*$ of (9a) and (9b), evidently we obtain the transformed values of length ${L_\text{U,E}}^*$ and of mass ${M_\text{U,E}}^*$. Moreover, when the square of the time of (8) is taken as the expression of the transformed cosmic time ${T_\text{U,E}}^{*2}$, then we obtain – *fully unexpected* – for ${M_\text{U,E}}^*$, ${L_\text{U,E}}^*$ and ${T_\text{U,E}}^{*2}$ one and the same number. This extremely spectacular observation results in the possibility to connect the transformed expressions of mass ${M_\text{U,E}}^*$, length ${L_\text{U,E}}^*$ and time ${T_\text{U,E}}^{*2}$ simply into one equation, given by

$${M_\text{U,E}}^* = {L_\text{U,E}}^* = {T_\text{U,E}}^{*2} = 4.03725(14)\times 10^{35}\,. \qquad (11)$$

It is evident that the spectacular non-dimensionality of (11) represents a particular state of unreality, which indicates the existence of an extraordinary state, seen in comparison to the in reality given basic units mass, length and time. It should be emphasized that this peculiar observation indeed reflects the experimental finding of the Klitzing-Effect (KE). The unexpected observation of the same number, resulting in (11), suggests the general validity of unification of basic units for the whole cosmos.

Besides, it should be emphasized that the disclosed possibility in this section 4 to describe each basic unit only by the same number can hardly be substantiated by the given physical argumentation in our times. Evidently, (11) is based on the proposed identity in section 2 of the category mass with frequency, i.e. with the category time. But, in an extensive manner valued, the reasoning of (11) can be indirectly supported by both the cosmological principle as well as by the mathematically unsolvable three-body-problem, which supports the DSS-model [4, 5], and which are generally valid. Moreover, a comprehensive analysis of (4) and (11) shows that the used constancy of the velocity of light c in the whole cosmos (reflecting the independence on the place of investigation) is only possible, when a universal validity of unification of the basic units is given. This statement is discussed in the next section.

5 The far-reaching findings about the category time

Based on all the data, shown in the preceding section 4, a comprehensive interpretation of the category time is possible by further analyzing the substance of (11). In fact, (8) together with (6a)–(6c), (9a)–(9c) and (11) yields a noticeable, far-reaching expression, given by

$$M_{\mathrm{U,E}} = T_{\mathrm{U,E}}\, T_{\mathrm{U,E}}^{*2} = T_{\mathrm{U,E}}\, M_{\mathrm{U,E}}^{*} \,. \qquad (12)$$

Considering (8), (9c) and (10a)–(10c), then (11) and (12), as well as the discussion of the results of the Hubble time $T_{\mathrm{U,E}}$ become more understandable for the interpretation of the essence of time, when we further postulate that the cosmic time $T_{\mathrm{U,E}}$ complies the remarkable "numerical" identity

$$T_{\mathrm{U,E}} = T_{\mathrm{U,E}}^{*} \,. \qquad (13)$$

The identity of the limit numbers of the real and unreal cosmic times, given by $T_{\mathrm{U,E}}$ and $T_{\mathrm{U,E}}^{*}$, can be confirmed, when analyzing (7) and (8) with respect to the results of vision and sound. According to (11) and (12), and especially to (13), the Hubble time $T_{\mathrm{U,E}}$ should be considered to be a particular magnitude, being numerically quite different to both $M_{\mathrm{U,E}}$ and $L_{\mathrm{U,E}}$ (the values of them are given in section 4) as well as to $M_{\mathrm{U,E}}^{*}$ and $L_{\mathrm{U,E}}^{*}$. Moreover, it must be pointed out that this specific exclusiveness of the cosmic time is also given in (11), where the real time $T_{\mathrm{U,E}}$, being in unreal state, appears in a "square". This important fact is a consequence of the specific circumstance that $T_{\mathrm{U,E}}$, in contrast to $M_{\mathrm{U,E}}^{*}$ and $L_{\mathrm{U,E}}^{*}$, is in this state used as a "real" magnitude, and that in form of $(1\,\mathrm{s})^{*}$. Thus, when we consider (12), it becomes evident that the circumstance of $T_{\mathrm{U,E}}$ can be in agreement with the DSS-model about the square relation of the 2D-state to the 1D-state "only", when the in (13) proposed numerical identity of the real and unreal cosmic time is given for the whole cosmos. It is obvious that this disclosure is confirmed by the existence of the Hubble effect [4].

Moreover, it should be pointed out that (11) and its general validity is based solely on the proposition of the validity of (4). It is given by the possibility to express the connection of the mass with the frequency, given in the reality related form by $f_{\mathrm{C}}(a_{\mathrm{G,E}})^{-1/2}$. The used rooty form of the gravitational value $a_{\mathrm{G,E}}$ is in full agreement with both the physical description of vision and sound [4, 5], and also the third law of Kepler, showing that the relation between the 2D-state and the 1D-state is given inevitably in squared form. This observation demonstrates the generally validity of the difference between the essence of frequency and time in the real state, which in the last consequence approves the correctness of (4) and (11).

All these discussed observations are important also for the confirmation of the general validity of (12), which can be obtained by comparison of this equation with the third law of Kepler. Due to the third law of Kepler, the orbital time of earth $t_{\mathrm{G,E}}$ is given by [4]

$$t_{\mathrm{G,E}} = (a_{\mathrm{G,E}})^{3/2}\, t_{\mathrm{E}} \,, \qquad (14)$$

where the specific time of earth t_{E} with respect to (1) is defined by

$$t_{\mathrm{E}} = \lambda_{\mathrm{G,E}}/c = 1.4797 \times 10^{-11}\ \mathrm{s} \,. \qquad (15)$$

Starting from (12), the application of it to the particular conditions of the earth yields the equation

$$M_{\mathrm{G,E}} = t_{\mathrm{E}}\, M_{\mathrm{U,E}}^{*} \,. \qquad (16)$$

Eqs. (12) and (16) demonstrate that the category time represents a connection of the real values of the masses $M_{\mathrm{U,E}}$ and $M_{\mathrm{G,E}}$ with the unreal value $M_{\mathrm{U,E}}^{*}$. Therefore, finally, we can draw the striking and for our investigation of the essence of time important conclusion that according to the effect of KE [1] and the DSS-model [4, 5], the category time suggests to be a magnitude to connect the 2D-state with the 1D-state, thus to be a factor of realization between reality and unreality.

As has been shown, this description of the essence of time is based on the proposed identity of the limit of mass with the limit of frequency, formulated by (4) and experimentally confirmed by many, quite different observations:

1. By the mysterious Klitzing-Effect (KE), which discloses the existence of an unresistance $R_{xx} = 0$ ohm at the Quantum-Hall-Effect (QHE),
2. by the physical description of the process of seeing and hearing [4, 5], resulting in the discovery of the DSS structure of space,
3. by the value of the gravitational number of the surface of the earth $a_{\mathrm{G,E}}$, determined from the mass and the radius of the earth, which modifies the limits of mass and frequency in (4),
4. by the results of the limiting current of the QHE [8, 9],
5. by the found approximate identity of the value of the Hubble time $T_{\mathrm{U,E}}$, experimentally determined by telescopes [4, 11], with the theoretically deduced limit time given by (8), which certifies the validity of the statements of (4) and (11),

6. by the coherence of the summarized result of the here listed experimental observations with the third law of Kepler [4], demonstrating that indeed the finding of the DSS-state and the possibility of unification of the basic units are in a perfect agreement with the assumed general validity of the third law of Kepler.

Considering all these experimental findings, the presented model of the essence of time should be viewed as physically confirmed.

6 Summarized conclusions

The physical description of vision and sound in form of the DSS-model [4,5], reflecting the experimental data of the Klitzing-Effect (KE) [1], reveals that the phenomenon time is given as a consequence of the interaction of the wave-related 2D-state with the 1D-state, i.e. the particular state, which shows exclusively real circumstances. This finding is the physical reasoning for the fact that the phenomenon time as well as the frequency in real form are always observable solely in connection with the category length [6] [9, pp. 9–10].

The analysis of the experimental data of basic units shows that surprisingly a generally valid identity of mass and frequency can be supposed, theoretically described by (4) and confirmed by different experimental data. These findings result in the possibility to formulate an equation, given by (11), in which, astonishingly, the transformed categories mass, length and time are given only by the same, single number, referring in the last consequence to the existence of an unreality in being. This observation has to be considered to be the physical description of the mysterious experimental effect of the $R_{xx} = 0$ ohm related Klitzing-Effect (KE), discovered in the Quantum-Hall-Effect (QHE) [1].

The third law of Kepler proves that the phenomenon time is observable only in the given state of the spatial three-dimensionality. This circumstance points out that – with respect to the DSS-model – the category time must be effective as a mediator for the coupling of the 2D-state with the 1D-state, thus being an important background of the realization of the category mass.

Summarizing all the presented experimental findings, it is allowed to conclude that the "unification" of the basic units length, mass, time and electric current appears to be a physical fact. Based on this important discovery, we can state that the phenomenon time is a dualistic factor, which is related on the one side to the localized, i.e. real 1D-state, noticeable as time, on the other side in the wavy, i.e. unreal 2D-state, in realistic form known as frequency. Evidently, this finding shows the existence of a substantial difference between time and frequency, and that to be a physical legitimate circumstance.

Acknowledgements

The author is indebted to *Prof. Walter Hansch*, University of Armed Forces, Muenchen, for the support of this subject and for his significant comments to some statements of this paper. He would also like to express special thanks to *Alexander Hirler* for his co-operation within the scope of critical reading of the paper and for its extensive completion for printable publication. Finally, the author expresses many thanks to *Klara Kirschner* for proof-reading of the English text.

Received on November 4, 2020

References

1. von Klitzing K., Dorda G. and Pepper M. New method for high-accuracy determination of fine-structure constant based on quantized Hall resistance. *Phys. Rev. Lett.*, 1980, v. 45, 494–497.
2. Hehl F. W. and Lämmerzahl C. Physical Dimensions/Units and Universal Constants: Their Invariance in Special and General Relativity. *Ann. Phys. (Berlin)*, 2019, 1800407.
3. Kose V. and Wöger W. Neuere empfohlene Werte von Fundamentalkonstanten. *Phys. Bl.*, 1987, v. 43, 397–399.
4. Dorda G, The Interpretation of the Hubble-Effect and of Human Vision Based on the Differentiated Structure of Space. *Prog. Phys.*, 2020, v. 16 (1), 3–9.
5. Dorda G. The Interpretation of Sound on the Basis of the Differentiated Structure of Three-Dimensional Space. *Prog. Phys.*, 2020, v. 16 (1), 15–19.
6. Smolin L. The Trouble with Physics. Houghton Mifflin Company, New York, 2006. In German: Die Zukunft der Physik. Deutsche Verlags-Anstalt, Muenchen, 2009, pp. 344–346.
7. Mende D. and Simon G. Physik. Gleichungen und Tabellen. Carl Hanser Verlag, Fachbuchverlag Leipzig, Muenchen, 2013.
8. Wittmann F. Magnetotransport am zweidimensionalen Elektronensystem von Silizium-MOS-Inversionsschichten. Dissertation, University of Armed Forces, Muenchen, Institute of Physics, Abb. 4–5.3, S. 143, 1992.
9. Dorda G. Quantisierte Zeit und die Vereinheitlichung von Gravitation und Elektromagnetismus. Cuvillier Verlag, Goettingen, ISBN 978-3-86955-240-8, 2010. In particular: pp 34–42, Fig. 2.1 and 2.2, and (2.14).
10. Dorda G. Die Struktur von Raum und Zeit und die Interpretation der Wärme. Cuvillier Verlag, Goettingen, ISBN 978-3-7369-9388-4, eISBN 978-3-7369-8388-5, 2016.
11. Bernhard H., Lindner K. and Schukowski M. Wissensspeicher Astronomie. Volk und Wissen Verlag, Berlin, ISBN 3-06-081705-7, 1995.

A Survey on Uncertainty Relations and Quantum Measurements:
Arguments for Lucrative Parsimony in Approaches of Matters

Spiridon Dumitru
(Retired) ex-Department of Physics, "Transilvania" University,
B-dul Eroilor 29, R-2200 Braşov, Romania,
E-mail: s.dumitru42@yahoo.com

This survey tries to investigate the truths and deficiencies of prevalent philosophy about Uncertainty Relations (UR) and Quantum Measurements (QMS). The respective philosophy, known as being eclipsed by unfinished controversies, is revealed to be grounded on six basic precepts. But one finds that all the respective precepts are discredited by insurmountable deficiencies. So, in regard to UR, the alluded philosophy discloses oneself to be an unjustified mythology. Then UR appear either as short-lived historical conventions or as simple and limited mathematical formulas, without any essential significance for physics. Such a finding reinforces the Dirac's prediction that UR "*in their present form will not survive in the physics of future*". The noted facets of UR motivate reconsiderations of associated debates on QMS. Mainly one reveals that, properly, UR have not any essential connection with genuine descriptions of QMS. For such descriptions, it is necessary that, mathematically, the quantum observables to be considered as random variables. The measuring scenarios with a single sampling, such are wave function collapse or Schrödinger's cat thought experiment, are revealed as being useless inventions. We propose to describe QMS as transmission processes for stochastic data. Note that, for existing quantum debates, the above UR–QMS revaluations, offer a few arguments for lucrative parsimony in approaches of matters. The unlucrative aspects of those debates have to be reconsidered too, probably in more or less speculative visions.

Motto 1: "I think one can make a safe guess that uncertainty relations in their present form will not survive in the physics of future."

P. A. M. Dirac, 1963

Motto 2: "The word 'measurement' has had such a damaging effect on the discussions that … it should be banned altogether in quantum mechanics."

J. S. Bell, 1990

Foreword

A. The present review-study germinates from some of our preceding more modest investigations some of them already published in this journal, Progress in Physics. Also, it was influenced by a number of opinions published by other scientists (opinions which, usually, are ignored in mainstream literature).

In the main, the study was stimulated by the known existence of numerous debates (unfinished controversies on unelucidated questions) regarding the foundations and interpretation of Quantum Mechanics (QM). The considered debates refer mainly to the significance of Uncertainty Relations (UR) and to the associated descriptions of Quantum Measurements (QMS). By their obstinate persistence, the mentioned debates delay and obstruct the desired (and expected) clarifications about some basic aspects of QM.

Within the here emerged text, we try to gather, systematize, improve, consolidate and mainly to present more argued our non-conventional viewpoints about the existing prevalent debates on UR, QMS and QM.

B. The here proposed article approaches step-by-step the following main items:

- i.1 A consistent Introduction which points out:
  (a) The nowadays existence of unfinished debates (disputes and controversies) about the meaning of UR and description of QMS;
  (b) The today necessity for search the truth about own philosophy of UR and description of QMS, regarded as relevant pieces for foundations/interpretation of QM;
- i.2 An inventory section which identifies the Basic Precepts of the prevalent philosophy regarding UR and QMS;
- i.3 A large section about most important deficiencies of the mentioned precepts. Within the respective section we concern on:
  (a) Detailed examinations of deficiencies specific to the respective precepts;
  (b) Elucidation, piece by piece, of the real value/meaning for each of the pointed out deficiencies;
- i.4 A first concluding section about the true significance of UR. In that section the current prevalent interpretation of UR is proved to be nothing but a veritable myth without any special or extraordinary value for physics. But such a proof reinforces the Dirac's prediction that UR

"in their present form will not survive in the physics of future";

i.5 A section containing considerations on description of QMS. The respective considerations are done in the light of the debates about deficiencies of dominant philosophy about UR. Also the measuring scenarios with a single sampling, such are wave function collapse or Schrödinger's cat thought experiment, are revealed as being superfluous fictions. We argue that the QMS descriptions should be approached additionally comparatively with the description and interpretation of UR. They must be discussed in new insights by regarding the measurements as transmission processes for stochastic data. (see our examples from Subsections 5.2 and 5.4 or from Appendices F and G);

i.6 A final section with some concluding remarks;

i.7 A supplementary section of Appendices containing:
(a) Technical/computational details — in seven cases, respectively;
(b) A copy of "A private letter from the late scientist J. S. Bell to the author".

C. Notes:

I. Through the elucidations referred to in item i.3 we offer genuine solutions for some controversial theoretical problems such are:
(a) The adequate form of UR for the supposed rebellious pairs of observables: L_z–φ (angular momentum — azimuthal angle), N–ϕ (number-phase) and E–t (energy-time);
(b) The case of macroscopic operators;
(c) The uniqueness (individuality) of Planck's constant;

II. In its essence, the suggested revaluation of UR and QMS philosophy does not disturb in any way the basic lucrative framework of usual QM (which keeps its known specific concepts, principles, theoretical models, computing rules and studied systems);
Moreover, I try to give arguments for lucrative parsimony in approaches of QM matters;
I believe that, to some extent, such a revaluation of UR-QMS prevalent philosophy can be beneficent for interpretation and understanding of QM. Potentially that revaluation can bring at least a modest contribution to non-conventional investigations of some open questions regarding views about UR, QMS, and QM.

III. My article tries to clarify certain past misunderstandings, of historical, philosophical, and cultural essence, which still persists in activities (of publishing and mainly of teaching nature), connected with QM;

IV. As a significant aspect, in my paper, the discussions are presented and detailed in forms accessible to readers with knowledge of QM at a not-advanced level. That is why in the version proposed here the article was conceived (especially through a number of detailed Appendices) as an accessible teaching material for those interested in QM education at undergraduate/graduate levels.

D. I think that, by its theme, style and writing level, my paper ensures the following desiderata:

- It approaches representative methodological and philosophical topics concerning the structure and the growth (interpretation and foundations) of QM investigated as a significant constituent of natural sciences;
- It can give a starting forum for the exchange of views and ideas among readers interested, in foundations of QM regarded as an important constituent of modern sciences;
- It identifies and highlights foundational issues, suggesting constructive and genuine solutions for approached problems;
- It offers a number of original opinions concerning some controversial theoretical/philosophical scientific problems;
- It initiates and develops discussions on the philosophy and epistemology of physics, at a level accessible to a wide class of readers (scientists, teachers and even students in physics, mathematics, chemistry or philosophy);
- It provides an argued appeal toward an increasing research field, namely to the one regarding the non-conventional approach of QM interpretation and foundations.

Given the above-mentioned aspects, I think that my article can offer a modest contribution to newly rising investigations on non-conventional views in quantum physics.

Braşov, November 26, 2020 *Spiridon Dumitru*

1 Introduction

Nearly a century until nowadays, in the publications regarding Quantum Mechanics (QM) and even other areas, have persisted discussions (debates and controversies) about the meaning of Uncertainty Relations (UR). Moreover UR in their entirety were ranked to a status of fundamental concept named Uncertainty Principle (UP) (for a bibliography of the better known specific publications see [1–12]). Mostly the respective discussions have credited UR/UP with considerable popularity and crucial importance, both in physics and in other domains. The mentioned importance was highlighted by compliments such as:

- UR are "*expression of the most important principle of the twentieth century Physics*" [13],
- UP is "*one of the cornerstones of quantum theory*" [9];

- UP "*epitomizes quantum physics, even in the eyes of the scientifically informed public*" [7].

But, as a fact, in spite of such compliments, in scientific literature of our days the essential aspects regarding UR/UP remain as unsolved and misleading questions. Today keeps their topicality many critiques reported during last decades, like the next ones:

- UR "*are probably the most controverted formula in the whole of the theoretical physics*" [14];
- "*Still now, 80 years after its inception, there is no general consensus over the scope and validity of this principle ('UP')*" [7];
- "*Overcoming the early misunderstanding and confusion, the concept* (notion of uncertainty — i.e. of UR/UP) "*grew continuously and still remains an active and fertile research field*" [8].

Note that the above reminded appreciations (compliments and critiques) regard mainly the own essence (intrinsic meaning) of UR/UP. But, within many texts about QM fundamentals, one finds also an adjacent topic which, historically, is a direct sub-sequence of the debates about the mentioned essence. The respective topic refers to the significance and description of Quantum Measurements (QMS).

Marked by the previously noted points, during the decades, the discussions about UR and QMS meaning and implications have generated a true prevalent philosophy (i.e. "*a group of theories and ideas related to the understanding of a particular subject*" [15]). For almost a century, the respective philosophy dominates in mainstream physics publications and thinking. It obstructs (delays) the expected progresses in clarifying some of main aspects regarding the fundamentals/interpretation of QM respectively the essentials of QMS problem. Add here the more alarming observation [16] that: "*there is still no consensus on … interpretation and limitations* of QM". Then it becomes of immediate interest to continue searches for finding the truth about own essence and consecutive topics of the UR/UP and QMS matters.

A search of the alluded type can be done (or facilitated at least) by a pertinent survey on deficiencies of the mentioned philosophy. Such a survey (of modest extent) we intend to present in this article. Our survey tries firstly to identify the basic elements of nowadays prevalent views within UR and QMS philosophy. Afterward we will investigate truth and value of the respective elements. Within the investigation we promote a number of re-considerations regarding the conventional (and now dominant) views about UR and QMS matters. Mainly we reveal the fact that the alluded views are discredited (and denied) by a whole class of insurmountable deficiencies, overlooked in the mainstream literature. So our survey aims to represent an unconventional analysis of the actual dominant philosophy about UR and QMS.The above-announced analysis germinated from some of our preceding investigations (see [17–21] and references). **Also, it was stimulated by a number of opinions due to other scientists (usually the respective opinions are ignored in dominant literature, but here they are highlighted by specifying the proper bibliographic sources)**. Through the present survey, we try to gather, extend, systematize, improve and consolidate the results of our mentioned investigations in order to present a more argued viewpoints about the approached topics.

In our survey, when it is usefully, we will appeal to the so called 'parsimony principle'(or 'law '). The respective principle (known also as Ockham's razor) will be applied as a heuristic method of simplicity which can be summarized [22] by the next two desiderata:

- "*Of two competing theories, the simpler explanation of an entity is to be preferred*".
- "*Entities are not to be multiplied beyond necessity*".

The mentioned principle will be accounted for in order that the text to be easy understood for readers (including students) not highly specialized.

By the present article-survey, through adequate arguments and details, we try to elucidate what is in fact the true meaning of UR, respectively to evaluate the genuine scientific aspects regarding QMS.

From the conclusions resulting from this survey the most important one is that, in its entirety, the actual prevalent philosophy about UR must be regarded as a veritable myth without any special or extraordinary status/significance for physics. This because, in reality, the UR reveal themselves to be nothing but short-lived historical conventions (in empirical, thought-experimental version) or simple and restricted formulas (in theoretical approach). But such a conclusion come in consonance, from another perspective, with the Dirac's guess [23] that: "*uncertainty relations in their present form will not survive in the physics of future*".

Add here the fact that, essentially, the above mentioned re-evaluation of UR and QMS philosophy does not disturb in any way the basic framework (principles, concepts, models and working rules) of usual QM. Furthermore, the QMS description remains as a distinct and additional subject comparatively with the elements of QM in itself. Add here the observation that, for existing quantum debates, the above UR–QMS revaluations give a few arguments for lucrative parsimony in approaches of matters. The unlucrative aspects of those debates have to be reconsidered, probably in more or less speculative visions.

The mentioned description of QMS requires to regard quantum observables* as true random variables. Also it must

---

***Drafting specifications:** (i) In the next parts of this article, for naming a physical quantity, we shall use the term "*observable*" (promoted by the UR and QMS philosophy literature), (ii) Also, according to the mainstream publications, we adopt the titles "*commuting*" or "*non-commuting*" observables for the QM quantities described by operators which "*commute*" respecbe dissociated of some fictive QMS scenarios with a unique sampling (such scenarios are schema with wave function collapse and Schrödinger's cat thought experiment). We recommend to describe QMS as transmission processes of stochastic data.

2 Basic precepts of UR–QMS prevalent philosophy

Firstly it must be pointed out the fact that, in spite of its prevalence inside of nowadays scientific debates, the actually dominant philosophy about UR and QMS germinates mainly from an old doctrine which can be called Conventional Interpretation of UR (CIUR). The mentioned doctrine (or dogma) was initiated by the Copenhagen School founders and, subsequently, during nine decades, it was promoted (or even extrapolated) by the direct as well indirect partisans (conformists) of the respective school. Currently CIUR enjoys of a considerable acceptance, primarily in QM studies but also in other thinking areas. Moreover, today, within the normative (mainstream/authoritarian) physics publications, CIUR dominates the leading debates about foundations and interpretation of QM.

But as a notable fact, in publications, CIUR doctrine, as well as most aspects of UR and QMS philosophy, are presented rather through independent or disparate assertions but not through a complete and systematized set of clearly defined "*precepts*" (considered as "*beliefs … accepted as authoritative by some group or school*" [24]). That is why, for a fruitful survey of the UR–QMS philosophy, it is of direct interest to identify such an set of ***B***asic ***P***recepts (***BP***) from which the mentioned assertions turn out to be derived or extrapolated. Note that the aforesaid set of precepts (i.e. the true core of CIUR doctrine along with prevalent philosophy of UR and QMS) can be collected by means of a careful examination of the today known publications. In its essence the respective collection can be presented as follows.

The history regarding Conventional Interpretation of UR (CIUR) began with two main generative elements which were the following ones:

(i) Heisenberg's "*Thought-Experimental*" (*TE*) relation:

$$\Delta_{TE}A \cdot \Delta_{TE}B \cong \hbar \quad or \quad \Delta_{TE}A \cdot \Delta_{TE}B \geqslant \hbar\,; \tag{1}$$

(ii) Robertson-Schrödinger relation of theoretical origin:

$$\Delta_{\Psi}A \cdot \Delta_{\Psi}B \geqslant \frac{1}{2}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle_{\Psi}\right|. \tag{2}$$

For introducing relation (1) in [25, 26] were imagined some "*Thought Experiments*" (*TE*) (or "gedanken" experiments). The respective *TE* referred on simultaneous measurements of two (canonically) conjugate observables A and B regarding a same quantum micro-particle. As such pairs of two observables were considered coordinate q and momentum p respectively time t and energy E. Then the quantities $\Delta_{TE}A$ and $\Delta_{TE}B$ were indicated as corresponding "*uncertainties*" of the imagined measurements, while $\hbar$ denotes the Planck's constant.

Relation (2) was introduced in [27, 28] and it is depicted as above in terms of traditional QM notations [29, 30]. The main features of the respective notations are reminded briefly below in Appendices A and B while some aspects regarding the Dirac's braket QM notations [29–32] are discussed in Appendix B.

Note here the fact that the right-hand side term from (2) is dependent on Planck's constant $\hbar$, e.g. $\left|\langle[\hat{A},\hat{B}]\rangle_{\Psi}\right| = \hbar$ when A and B are (canonically) conjugate.

Starting from the generative elements (1) and (2), CIUR doctrine jointly with UR and QMS philosophy have been evolved around the following ***B***asic ***P***recepts (***BP***):

- ***BP*₁**: Quantities $\Delta_{TE}A$ and $\Delta_{\Psi}A$ from relations (1) and (2), have similar significances of measuring uncertainties for the observable A. Consequently, the respective relations should be regarded as having a same meaning of Uncertainty Relations (UR) concerning the simultaneous measurements of observables A and B. Such a regard is fortified much more by the fact that
$$\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle_{\Psi}\right| = \hbar$$
when A and B are (canonically) conjugate.
- ***BP*₂**: In case of a solitary observable A, for a micro-particle, the quantities $\Delta_{TE}A$ or $\Delta_{\Psi}A$ can have always an unbounded small value. Therefore such an observable should be considered as measurable without any uncertainty in all cases of micro-particles (systems) and states.
- ***BP*₃**: For two commuting observables A and B (whose operators $\hat{A}$ and $\hat{B}$ commute, i.e. $[\hat{A},\hat{B}] = 0$) relation (2) allows for the product $\Delta_{\Psi}A \cdot \Delta_{\Psi}B$ to be no matter how small. Consequently the quantities $\Delta_{\Psi}A$ and $\Delta_{\Psi}B$ can be unlimited small at the same time. Such observables have to be regarded as being compatible, i.e. measurable simultaneously and without interconnected uncertainties, for any micro-particle (system) or state.
- ***BP*₄**: In case of two non-commuting observables A and B (described by operators $\hat{A}$ and $\hat{B}$ which do not commute, i.e. $[\hat{A},\hat{B}] \neq 0$) the relation (2) shows that the product $\Delta_{\Psi}A \cdot \Delta_{\Psi}B$ has as lower bound a non-null and $\hbar$-dependent quantity. Then the quantities $\Delta_{\Psi}A$ and $\Delta_{\Psi}B$ can be never reduced concomitantly to null values. For that reason the respective observables must be accounted as measurable simultaneously only with non-null and interconnected uncertainties, for any situation (particle/state). Viewed in a pair such observables are proclaimed as being incompatible, respectively complementary when they are (canonically) conjugate.

---

tively "*do not commute*", (iii) For improving fluency of our text some of the corresponding mathematical notations, formulas and proofs are summarized briefly and unitary in few Appendices located in the final of the article.

- ***BP*$_5$**: The main elements of CIUR doctrine and UR philosophy show quantum particularities of uniqueness comparatively with other non-quantum areas of physics. Such elements are the very existence of relations(1) and (2), the above asserted measuring features and the discriminating presence of the Planck's constant $\hbar$.
- ***BP*$_6$**: For glorifying the precepts ***BP*$_1$–*BP*$_5$** and adopting the usages of dominant literature, UR philosophy in its entirety should be ranked to a status of fundamental concept named Uncertainty Principle (UP).

Add here the observation that, in their wholeness, CIUR doctrine conjointly with UR and QMS prevalent philosophy emerge completely from the assertions embedded in basic precepts "***BP*$_1$–"***BP*$_6$**.

3 *D*eficiencies (*D*) of the mentioned precepts

The above mentioned emergence conceals a less popularized fact namely that each of the precepts ***BP*$_1$–*BP*$_6$** is discredited (and denied) by insurmountable deficiencies. Such a fact can be revealed through a deep analysis of the respective precepts, an analysis which is of major importance for an authentic and fruitful survey of UR and QMS prevalent philosophy. That is why here below we aim to reveal the most significant ones of the mentioned deficiencies. They will be presented in a meaningful ensemble, able to give an edifying global appreciation regarding the mentioned philosophy. The referred ensemble includes as distinct pieces the following ***D***eficiencies (***D***):

3.1 D_1: Provisional character of relation (1)

Now it must be noted firstly the aspect that, through an analysis of its origins, relation (1) shows only a provisional (transient) character. This because it was founded [25, 26] on old resolution criterion from optics (introduced by Abe and Rayleigh — see [33]). But the respective criterion was surpassed through the so-called super-resolution techniques worked out in modern experimental physics (see [34–38] and references). Then by means of of the mentioned techniques can be imagined some interesting "*Super-Resolution-Thought-Experiments*" (*SRTE*). Through such *SRTE* for two (canonically) conjugate observables A and B, instead of *TE*-uncertainties $\Delta_{TE}A$ and $\Delta_{TE}B$ from (1), it becomes possible to discuss situations with some *SRTE*-uncertainties denoted as $\Delta_{SRTE}A$ and $\Delta_{SRTE}B$. For the respective *SRTE*-uncertainties, instead of Heisenberg's restrictive formula (1) (first version), can be suggested some CIUR-discordant relations like as

$$\Delta_{SRTE}A \cdot \Delta_{SRTE}B < \hbar. \tag{3}$$

Note that an experimental example of discordant relation of (3)-type was mentioned in [39] (where the UR (1) "*would be violated by close to two orders of magnitude*").

Now one observes that, from the our days scientific perspective, *SRTE* relations like (3) are suitable to replace the old Heisenberg's formula (1) (second version). But such suitability invalidates a good part of the precept ***BP*$_1$** and, additionally, it incriminates the CIUR doctrine and UR–QMS philosophy in connection with one of their main (generative) element.

It is surprising that, after invention of the super-resolution techniques, the mainstream (normative /authoritarian) publications connected with UR–QMS philosophy avoided a just and detailed evaluation of the respective techniques. Particularly, even after eight year after the result reported in [39], almost all of the dominant publications omit to discuss the respective result. The surprise is evidenced to a great extent by the fact that parsimony desiderata noted in Section 1 offer a viable argumentation for completing the evaluations and discussions oft the mentioned kind.

Another infringement (violation) of Heisenberg's relation (1) was reported in [40] as an experimental result. That report is criticized vehemently by CIUR partisans [12]. The respective criticism is done in terms of a few un-argued (and un-explained) accusatory-sentences. But it is expected that, if they are justifiable, such kind of critiques should be grounded on precise technical details and arguments. This in order that they to be credible.

Curiously is also the fact that, over the past decades within the UR philosophy, the debates have neglected the older criticisms of the relation (1) due to K. Popper [41].

Taking into account the above revealed aspects one can say that the precept ***BP*$_1$** proves oneself to be a misleading (even harmful) basic element for CIUR doctrine and UR–QMS philosophy. But such a proof is a first argument for reporting that the respective doctrine and philosophy cannot be accepted as solid (and credible) scientific constructions.

3.2 D_2: Significance of quantities from relation (2)

The term "*uncertainty*" used within CIUR doctrine for quantities $\Delta_\Psi A$ and $\Delta_\Psi B$ from (2) is groundlessly because of the following considerations. According the theoretical framework of QM, by their definitions, the respective quantities signify genuinely the standard deviations of the observables A and B regarded as random variables (see below Appendix A). With such significances the alluded quantities refer to intrinsic (own) properties (known as fluctuations) of the considered particle but not to characteristics of the measurements performed on respective particle. In fact, on a one hand, for a measured particle in a given state (described by certain wave function Ψ) the quantities $\Delta_\Psi A$ and $\Delta_\Psi B$ have unique and well definite values. On the other hand for the same particle/state the measuring uncertainties regarding the observables A and B can be changed through the improvements or deterioration of experimental devices/techniques.

The above revealed QM significances for quantities $\Delta_\Psi A$ and $\Delta_\Psi B$ are genuinely preferable comparatively with the assertions from the precepts ***BP*$_1$–*BP*$_4$** promoted by CIUR doctrine and UR–QMS philosophy. But such a preference is completely congruent with the previously mentioned desiderata of parsimony principle.

3.3 D_3: Limitations of relation (2)

Relation (2) has only limited validity within the complete theoretical framework of QM. This because, as it is detailed below in Appendix A, for observables A and B, relation (2) is only a restricted consequence of the generally valid Cauchy-Schwarz formula, given in (A.2). From such a general formula the relation (2) results iff (if and only if) in circumstances when the conditions (A.3) are satisfied. In the respective circumstances in addition to relation (2)/ (A.7) from (A.2) arises also the formula (A.6). It is worthy to note that the mentioned particularities regarding the validity of the relation (2) discredit indirectly the precept $\boldsymbol{BP_1}$ of CIUR doctrine and UR–QMS philosophy. In their essence the specifications recorded here are nothing but concretizations of parsimony desiderata regarding the respective doctrine and philosophy.

3.4 D_4: On solitary observables

It is surprising to find that, within UR–QMS philosophy debates, the problem of solitary observables is not discussed carefully. Particularly, were neglected discussions regarding the measurements of such observables. This although the respective discussions can be sub-summed to the question of simultaneous measurements of two observables. Such a sub-summation can be imagined by means of the Thought Experiments (*TE*) which motivated the conventional relation (1). Namely, for example, if in the respective *TE* it is of interest only the quantity $\Delta_{TE}A$, by ignoring completely the quantity $\Delta_{TE}B$, one can say that $\Delta_{TE}A$ can be unlimited small. Therefore the observable A, regarded as a solitary variable, appears as measurable without any uncertainty in all cases. But, on the other hand, if the same solitary observable A is analyzed in terms of relation (2), it cannot be associated with an unlimited small value for the quantity $\Delta_\Psi A$. This because, form a QM perspective, $\Delta_\Psi A$ has a unique and well definite value, evaluated through the corresponding wave function Ψ. Consequently, even in the cases of solitary observables, the CIUR doctrine and the UR–QMS philosophy cannot provide a clear and unequivocal approach as it is suggested by precept $\boldsymbol{BP_2}$.

3.5 D_5: About commutable observables

According to the precept $\boldsymbol{BP_3}$ for two observables A and B, whose associated operators $\hat{A}$ and $\hat{B}$ are commutable, relation (2), allows for the product $\Delta_\Psi A \cdot \Delta_\Psi B$ to be however small. Then the quantities $\Delta_\Psi A$ and $\Delta_\Psi B$ can be unlimited small at the same time. Such observables are supposed compatible, they being measurable simultaneously and without interconnected uncertainties for any micro-particle (system) or state.

But, as it was shown above in deficiency $\boldsymbol{D_2}$, the mentioned assertions from $\boldsymbol{BP_3}$, conflict with the genuine significance of the quantities $\Delta_\Psi A$ and $\Delta_\Psi B$. This because both $\Delta_\Psi A$ and $\Delta_\Psi B$ have unique values, determined theoretically by the wave function Ψ which describe the considered state of particle. Or it is possible to have some "rebellious situations" in which the respective values of $\Delta_\Psi A$ and $\Delta_\Psi B$ to be simultaneously non-zero but finite entities, even the corresponding observables are commutable.

Such a "rebellious situation" can be found [20] for the observables P_x and P_y (Cartesian moments) regarding a micro-particle situated in a potential well of a rectangular 2D configuration. If the well walls are inclined towards the X and Y axes, the both the quantities $\Delta_\Psi P_x$ and $\Delta_\Psi P_y$ have non-zero but finite values. In that situation for P_x and P_y, besides the relation (2), it is satisfied however the formula (A.2) with $\left|(\delta_\Psi \hat{P}_x \Psi, \delta_\Psi \hat{P}_x \Psi)\right|$ as a non-null quantity.

The above remarks show that, in fact, the cases of commutable observables require to repudiate firmly the precept $\boldsymbol{BP_3}$. Additionally we think that the same cases should be regarded in the spirit of parsimony principle desiderata, by their consideration in QM terms reminded briefly in Appendices A and B.

3.6 D_6: Cases of angular observables L_z and φ

The precept $\boldsymbol{BP_4}$ stipulates that, as a principle, two non-commutable observables A and B cannot be measured simultaneously because the product $\Delta_\Psi A \cdot \Delta_\Psi B$ has a non-null lower bound. But the respective stipulation is contradicted by some rebellious pairs of observables. Such a pair, widely discussed, is $L_z - \varphi$ (angular momentum — azimuthal angle), regarded in certain particular situations. The respective contradiction was probably the most inciting subject of debates during the history of CIUR doctrine and UR–QMS philosophy (see [5, 17–20, 42–55]). The mentioned debates regarded mainly the quantum rotations which can be called "*L_z-non-degenerate — circular — rotations*" (*L_z-ndcr*). But, besides of that situations, in QM framework can be discussed also other kinds of rotations, of direct significance for $L_z - \varphi$ pair. Such kinds are the ones regarding the rotational eigenstates of a Quantum Torsion Pendulum (QTP) and respectively the "*L_z-degenerate — spatial — rotations*" (*L_z-dsr*). The true situations of the L_z–φ pair in relation with all kinds of the mentioned rotations will be discussed below in more details.

3.6.1 D_{6a}: About non-degenerate circular rotations

Let us discuss now the cases of *L_z-non-degenerate — circular — rotations* (*L_z-ndcr*). As systems of with *L_z-ndcr* can be quoted the following ones: (i) a particle (bead) on a circle, (ii) an 1D rotator and (iii) non-degenerate spatial rotations of a particle on a sphere or of an electron in a hydrogen atom respectively. The mentioned spatial rotations are considered as L_z-non-degenerate if the magnetic quantum number m (associated with L_z) has a unique value (while, of course, all other specific (orbital) quantum numbers have well-defined values).

The rotations of respective systems are described through the wave functions given by

$$\Psi(\varphi) = \Psi_m(\varphi) = (2\pi)^{-\frac{1}{2}} \cdot \exp(im\varphi). \tag{4}$$

Here φ is an ordinary polar coordinate (angle) with the corresponding mathematical characteristics [56] i.e. $\varphi \in [0, 2\pi)$ and number m gets only one value from the set $m = 0, \pm 1, \pm 2, \dots$. Also in (4) the wave function $\Psi(\varphi) = \Psi_m(\varphi)$ has the property $\Psi(0) = \Psi_m(2\pi - 0) := \lim\limits_{\varphi \to 2\pi - 0} \Psi_m(\varphi)$.

In the same context, according to the known QM framework [29], L_z and φ should be regarded as polar observables, described by the conjugated operators and commutator represented as follows

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \varphi}, \quad \hat{\varphi} = \varphi\cdot, \quad \left[\hat{L}_z, \hat{\varphi}\right] = -i\hbar. \tag{5}$$

Therefore the conventional relation (2) motivates as a direct consequence the next formula

$$\Delta_\Psi L_z \cdot \Delta_\Psi \varphi \geqslant \frac{\hbar}{2}. \tag{6}$$

Now it is easy to observe that this last formula is explicitly inapplicable in cases described by wave functions (4). This because in such cases, for the quantities $\Delta_\Psi L_z$ and $\Delta_\Psi \varphi$ associated with the pair L_z–φ, one obtains the following values

$$\Delta_\Psi L_z = 0, \quad \Delta_\Psi \varphi = \pi \cdot (3)^{-\frac{1}{2}}. \tag{7}$$

But such values for $\Delta_\Psi L_z$ and $\Delta_\Psi \varphi$ are evidently incompatible with the conventional relation (2)/(6).

In order to avoid the above revealed incompatibility in many mainstream publications the CIUR partisans promoted some unusual ideas such are:

- For L_z and φ operators and commutator, instead of current expressions (5), it is conveniently to adopt other new denotations (definitions).
- The formula (6) must be abandoned/proscribed and replaced by one (or more)"*modified L_z– φ UR*" able to mime the conventional relation (2) for the L_z– φ pair.

The alluded ideas were promoted through the conception of "*impossibility of distinguishing … between two states of angle differing by 2π*". But such a conception has not any realistic sense in cases of circular rotations. This because in such cases the angle φ has as physical range the interval $[0, 2\pi)$. Moreover in the respective cases the wave functions (4) are normalized on the same interval but not on other strange domains.

As regards the "*modified L_z– φ UR*", along the years, by means of some circumstantial (and more or less fictitious) considerations, were proposed a lot of such relations. In terms of usual QM notations (summarized below in Appendix A), the alluded "*modified L_z– φ UR*" can be written generically as follows

$$f(\Delta_\Psi L, \Delta_\Psi g(\varphi)) \geqslant \hbar \cdot \langle s(\varphi) \rangle_\Psi. \tag{8}$$

Here $f(\Delta_\Psi L, \Delta_\Psi g(\varphi))$, $g(\varphi)$ and $s(\varphi)$ denote some specially invented functions depending on the corresponding arguments. Note that some of the mostly known concrete examples of relations (8) can be found collected in [55].

Now it should be noted the fact that the "*modified L_z– φ UR*" such are (8) show some troubling features like the following ones:

- Regarded comparatively, the mentioned "*modified L_z– φ UR*" are not mutually equivalent. This despite of the fact that they were invented in order to substitute the same proscribed formula (6). Consequently, none of that modified relations, is agreed unanimously as a suitable model able to give such a substitution.
- Relations (8) are in fact ad hoc artifices without any source in mathematical framework of QM. Then, if one wants to preserve QM as a unitary theory, like it is accredited in our days, the relations (8) must be regarded as unconvincing and inconvenient (or even prejudicial) inventions.
- In fact in relations (8) the relevant angular quantities $\Delta_\Psi L_z$ and $\Delta_\Psi \varphi$ are substituted more or less factitious with the adjusting functions $f(\Delta_\Psi L_z, \Delta_\Psi g(\varphi))$, $g(\varphi)$ and $s(\varphi)$. But, from a genuine perspective, such substitutions, and consequently the corresponding relations, are only mathematical constructs but not elements with useful physical significance. Of course that such constructs overload (or even impede) the scientific discussions by additions of extraneous entities which are not associated with true information about the real world.

Then, for a correct evaluation of the facts, all the aspects regarding relations (8) versus (6) ought to be judged by taking into consideration the parsimony principle desideratum: "*Entities are not to be multiplied beyond necessity*". Such an evaluation can be started by clarifying firstly the origin and validity conditions of the formula (6) regarded as descendant of conventional relation (2). For the respective clarification it is usefully to see some QM elements briefly summarized in Appendix A.

So it can be observed easy that, in its essence, the relations (2) follow from the generally valid formulas (A.2) pertaining to the mathematical framework of QM. But, attention, (2) results correctly from (A.2) iff (if and only if) when it is satisfied the condition (A.3). In other cases (2) are not valid at all. Such an invalidity is completely specific for the cases of L_z– φ pair in relations with situations described by the wave functions (4). This because in respective cases instead of conditions (A.3) it is true the relation

$$\left(\hat{L}_z \Psi, \hat{\varphi}\, \Psi\right) = \left(\Psi, \hat{L}_z\, \hat{\varphi}\, \Psi\right) + i\hbar. \tag{9}$$

Therefore, for systems described by the wave functions (4), the formula (6) is invalid by its essence.

Now note that, even when the condition (A.3) is not satisfied, according to the QM general formula (A.2), for the discussed situations it is true the relation

$$\Delta_\Psi L_z \cdot \Delta_\Psi \varphi \geqslant \left| \left(\delta_\Psi \hat{L}_z \Psi,\ \delta_\Psi\, \hat{\varphi}\, \Psi \right) \right| \tag{10}$$

written in compliance with definitions (4) and (5). But, attention, in respective situations the last relation (10) degenerates into trivial equality "*0=0*". Add here the fact that relation (10) is completely equivalent with the formula (C.13) deductible within Fourier analysis.

The above presented details argue undoubtedly the view that in cases with L_z *-ndcr* the L_z– φ pair must to satisfy not the troublesome formula (6) but the QM justified relation (10) (which in fact reduces itself to banal equality "*0=0*"). Such an argued view clarifies all disputes regarding the mentioned cases. Moreover the same view disproves the idea of some "*entities … multiplied beyond necessity*" (such are the modified UR (8)) intended to replace the inoperative relation (6).

3.6.2 D_{6b}: Case of Quantum Torsion Pendulum (QTP)

The case of Quantum Torsion Pendulum (QTP) regards a quantum harmonic oscillator with torsional rotations [19, 20, 55]. Such an oscillator can be considered as the simplest theoretical model for molecular twisting motion ("*change in the angle between the planes of two groups of atoms*" [57]). For a QTP oscillating around the z-axis the Hamiltonian operator has the form

$$\hat{H} = \frac{1}{2I}\,\hat{L}_z^2 + \frac{1}{2}\,I\omega_0^2\hat{\varphi}^2. \tag{11}$$

Here φ denotes the twisting angle with domain $\varphi \in (-\infty, +\infty)$ while the operators $\hat{L}_z$ and $\hat{\varphi}$ obey the rules (5). The other symbols from (11) are: I and ω_0 represent the momentum of inertia respectively the (undamped) resonant frequency ($\omega_0 = \sqrt{\kappa/I}$ while κ = torsion elastic modulus).

By means of Schrödinger equation $E\Psi = \hat{H}\Psi$ one finds that the QTP eigenstates are described by the wave functions

$$\Psi_n(\varphi) = \Psi_n(\xi) \propto \exp\left(-\frac{\xi^2}{2}\right) \cdot \mathcal{H}_n(\xi), \quad \xi = \varphi\sqrt{\frac{I\omega_0}{\hbar}}. \tag{12}$$

These wave functions correspond to the oscillation quantum numbers $n = 0, 1, 2, 3, \ldots$ and energy eigenvalues $E_n = \hbar\omega_0\left(n + \frac{1}{2}\right)$. In (12) $\mathcal{H}_n(\xi)$ represent the Hermite polynomials of ξ.

For each of the states (12) for observables L_z and φ associated with the operators (5) one obtains the expressions

$$\Delta\varphi = \sqrt{\frac{\hbar}{I\omega_0}\left(n+\frac{1}{2}\right)}, \qquad \Delta L_z = \sqrt{\hbar I\omega_0\left(n+\frac{1}{2}\right)},$$
$$\left|\left(\Psi, \left[\hat{L}_z, \hat{\varphi}\right]\right)\right| = \hbar, \tag{13}$$
$$\Delta\varphi \cdot \Delta L_z = \hbar \cdot \left(n + \frac{1}{2}\right).$$

These expressions show the fact that, for each QTP eigenstate, the L_z– φ pair satisfies the relation (6)/(2). But note that the respective fact is due to the circumstance that in the mentioned case, in relation with the wave functions (12), the operators $\hat{L}_z$ and $\hat{\varphi}$ satisfy a condition of (A.3) type, i.e. $(\hat{L}_z\Psi, \hat{\varphi}\,\Psi) = (\Psi, \hat{L}_z\,\hat{\varphi}\,\Psi)$.

3.6.3 D_{6c}: On degenerate spatial rotations

Let us now regard the cases of L_z *–degenerate-spatial-rotations* (L_z*-dsr*). Such kinds of rotations refer [20, 21, 55] to states of: (i) a particle on a sphere, (ii) a 2D rotator and (iii) an electron in a hydrogen atom. The respective rotations are L_z-degenerate in sense that the magnetic quantum number m (associated with L_z) has multiple values while the other quantum numbers have unique values. A particle on a sphere or a 2D rotator are in a L_z *-dsr* when the orbital number l has a unique value greater than zero while m can take all the values $m \in [-l, +l]$. Then the corresponding rotations are described through the global wave function

$$\Psi(\varphi) = \Psi_l(\vartheta, \varphi) = \sum_{m=-l}^{m=+l} c_m \cdot Y_{lm}(\vartheta, \varphi). \tag{14}$$

Here ϑ and φ denote polar respectively azimuthal angles with $\vartheta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$. In (14) $Y_{lm}(\vartheta, \varphi)$ denote spherical functions while c_m are coefficients normalized through the condition $\sum_{m=-l}^{m=+l} |c_m|^2 = 1$. Also the wave functions $\Psi_l(\vartheta, \varphi)$ from (14) have the property $\Psi_l(\vartheta, 0) = \Psi_l(\vartheta, 2\pi - 0) := \lim_{\varphi\to 2\pi-0} \Psi_l(\vartheta, \varphi)$. In a direct connection with such a property the operators $\hat{L}_z$ and $\hat{\varphi}$ obey the rules (5).

Now let us regard what are the peculiarities of the L_z- *dsr* cases in respect with the controversial relation (6). Principled, such a regard demands that, by using the formulas (5) and (14), to evaluate the corresponding expressions for the quantities $\Delta_\Psi L_z$, $\Delta_\Psi\varphi$ and $\left|(\Psi, [\hat{L}_z, \hat{\varphi}\,]\Psi)\right|$. With the respective expressions one finds possibilities that the relation (6) to be or not to be satisfied. Of course that such possibilities are conditioned by the concrete values of the coefficients c_m. But note that, if the relation (6) is not satisfied, the fact appears because essentially in such a situation the condition (A.3) is not fulfilled. Add here the important observation that, independently of validity for relation (6), in all cases of L_z *-dsr* the L_z– φ pair obeys the prime QM relation (A.2) through adequate values for the quantities $\Delta_\Psi L_z$, $\Delta_\Psi\varphi$ and $\left|(\delta_\Psi\hat{L}_z\Psi, \delta_\Psi\hat{\varphi}\,\Psi)\right|$. The previous considerations offer a clear evaluation of the situation for L_z- *dsr* cases relatively to the conventional relation (2) and precept BP_4.

Summing up of deficiencies $\boldsymbol{D_6}$ (including $\boldsymbol{D_{6a}}$, $\boldsymbol{D_{6b}}$ and $\boldsymbol{D_{6c}}$): The above discussion about the three kinds of rotations reveals the deficiencies of the conventional relation (2) and of the associated precept $\boldsymbol{BP_4}$ in regard with the non-commutable observables L_z and φ. But such revealing is nothing but a direct and irrefutable incrimination of CIUR doctrine and UR–QMS philosophy.

3.7 D_7: On number and phase observables

The pair N and ϕ (number and phase) is another couple of rebellious non-commutable observables which contradict the corresponding stipulation from the precept $\boldsymbol{BP_4}$ of UR–QMS philosophy. That contradiction emerged in connection with the associated operators $\hat{N}$ and $\hat{\phi}$. The respective operators were introduced by means of the ladder (lowering and raising) operators $\hat{a}$ and $\hat{a}^+$, destined to convert some QM calculations procedures from an analytical version to an algebraic one. Through the respective connection, by taking as base the relation $[\hat{a}, \hat{a}^+] = 1$, it was inferred the commutation formula $[\hat{N}, \hat{\phi}] = i$.

The last noted formula motivated the idea that operators $\hat{N}$ and $\hat{\phi}$ must satisfy the conventional relation (2) with both $\Delta_\Psi N$ and $\Delta_\Psi \phi$ as non-null quantities. But afterward it was found the fact that, in the case of a harmonic oscillator eigenstates, one obtains $\Delta_\Psi N = 0$ and $\Delta_\Psi \phi = \pi \cdot (3)^{-\frac{1}{2}}$ i.e. a violation of the relation (2). Of course that such a fact leads to a deadlock for harmonization of N–ϕ observables with the CIUR doctrine and UR–QMS philosophy. Note that this deadlock is completely analogous with the one regarding to L_z–φ observables in the above discussed case of L_z-*ndcr* (L_z-non-degenerate — circular — rotations).

For avoiding the mentioned N–ϕ deadlock in many publications were promoted various adjustments (see [6, 43, 48, 58–61] and references therein). But it is easy to observe that the respective adjustments regarded the conventional relation (2) as an absolute mark and tried to adapt accordingly the pair N–ϕ for a description of a harmonic oscillator. So it was suggested to replace the original operators $\hat{N}$–$\hat{\phi}$ by some ad hoc "*adjusted*" (*adj*) operators $\hat{N}_{adj}$ and $\hat{\phi}_{adj}$, able to generate formulas resembling (more or less) with the conventional relation (2) (examples of such adjusted operators can be found in the literature of recent decades). However it is very doubtfully that the corresponding "*adjusted observables*" N_{adj} and ϕ_{adj} can have natural (or even useful) physical significances. Moreover, until now, it not exist a unanimously agreed conception able to guarantee a true elucidation regarding the status of number-phase observables relatively to terms of CIUR doctrine and UR philosophy.

Our opinion is that a genuine clarification of the N–ϕ problem can be done similarly with the above discussed situation of L_z–φ observables in the cases of L_z-*ndcr*. More exactly we have to note that the disagreement of N–ϕ pair with the conventional relation (2) results from fact that in such a case the respective relation is mathematically incorrect. The aforesaid incorrectness is due mainly to the circumstance that, in cases of a linear oscillator eigenstates, the N–ϕ pair does not satisfy the essential condition (A.3). This because in that cases for the operators $\hat{N}$–$\hat{\phi}$ is true the formula $(\hat{N}\Psi, \hat{\phi}\Psi) = (\Psi, \hat{N}\hat{\phi}\Psi) + i$ which evidently infringes the condition (A.3). But it should be pointed out that, even in the mentioned cases, the $\hat{N}$–$\hat{\phi}$ operators satisfy the primary relation (A.2) which degenerates into trivial equality "*0 = 0*".

We think that the above noted opinion gives a natural and incontestable solution for the problem regarding the N–ϕ pair versus the conventional relation (2). Accordingly the fictional operators $\hat{N}_{adj}$ and $\hat{\phi}_{adj}$, of an ad hoc adjusted essence, proves themselves to be nothing but "*entities … multiplied beyond necessity*".

So it can be said that the situation of observables N and ϕ contradict directly the precept $\boldsymbol{BP_4}$ in connection with non-commutable observables. Consequently, the respective situation invalidates completely one of basic elements of CIUR doctrine and UR–QMS philosophy.

3.8 D_8: Concerning the energy — time pair

Closely to the conventional views of CIUR doctrine and UR–QMS philosophy the pair of observables E–t (energy-time) was subject for a large number of controversial discussions (e.g. in works [5, 6, 62–64], in their references and, certainly, in many other publications). The alluded discussions were generated within following circumstances. On one hand, accordingly to the mentioned views, E and t are regarded as conjugated observables, having to be described by the next operators and commutator

$$\hat{E} = i\hbar\frac{\partial}{\partial t}, \qquad \hat{t} = t\cdot, \qquad \left[\hat{E}, \hat{t}\right] = i\hbar. \tag{15}$$

Then the operators $\hat{E}$ and $\hat{t}$ should satisfy the conventional relation (2) in a nontrivial version. On the other hand, because of the fact that, in terms of usual QM, the time t is a deterministic but not random variable, for any quantum situation one finds the following expressions $\Delta_\Psi E$ = "*a finite quantity*" respectively $\Delta_\Psi t \equiv 0$. But these expressions invalidate the relation (2) and consequently the E–t pair shows an anomaly in respect with the alluded conventional ideas, especially with the precept $\boldsymbol{BP_4}$. For avoiding the noted anomaly, within the literature about E–t pair, it was substituted the unsuitable relation (2) by some adjusted formulas written generically as follows

$$\Xi\, E \cdot \Xi\, t \geqslant \frac{\hbar}{2}. \tag{16}$$

The so introduced quantities ΞE and Ξt have various significances such are: (i) line-breadth and half-life of a decaying excited state, (ii) frequency domain and temporal widths of a wave packet, (iii) $\Xi E = \Delta_\Psi E$ and $\Xi\, t = \Delta_\Psi A \cdot ({}^{d\langle A\rangle}\!/_{dt})^{-1}$, with A = an arbitrary observable.

As regards the adjusted formulas (16) note firstly the fact that various of their versions are not congruent with the original conception of relation (2). Also the respective versions are not mutually equivalent from a mathematical (theoretical) viewpoint. So they have no reasonable justification in the true QM framework. Moreover in specific literature none of the formulas (16) is accepted unanimously as a correct (or natural) substitute for conventional relation (2).

Now it is the place to present the following clarifying remarks. Even if the *E–t* pair is considered to be described by the operators (15), according to the true QM terms, one finds the relation

$$\left(\hat{E}\Psi, \hat{t}\Psi\right) = \left(\Psi, \hat{E}\,\hat{t}\Psi\right) - i\hbar. \tag{17}$$

By comparing this relation with condition (A.3) one sees directly that the *E-t* pair cannot ever satisfy the respective condition. This is the essential reason because of which for the *E–t* pair the conventional relation (2) is not applicable at all. Nevertheless, for the same pair described by the operators (15), the QM relation (A.2) is always true. But because in QM the time *t* is a deterministic (i.e. non-stochastic) variable in all cases the respective true relation degenerates into the trivial equality "*0 = 0*".

The above noted comments lead to the next findings:

- In case of the *E–t* pair the conventional views (of CIUR doctrine and UR–QMS philosophy) are completely nonfunctional.
- Genuinely, within a true QM framework, the time *t* is in fact a pure deterministic (non-stochastic) quantity without any standard deviation (or fluctuation).

But, taken together, such findings about time-energy pair must be reported as a serious and insurmountable deficiency of CIUR doctrine and UR–QMS philosophy.

3.9 D_9: Atypical analogues of UR (1) and (2)

By basic precept $\mathbf{BP}_5$ the UR philosophy claims idea that relations (1) and (2) possess an essential typicality represented by their QM uniqueness related with the systems of atomic size. Consequently, the respective relations should not have analogues in other areas of physics or for systems of radically different sizes. But the respective idea is definitely denied by some example that we will present below.

3.9.1 D_{9a}: Classical Rayleigh formula

As a first example of an atypical analogue of the UR (1) can be quoted the formula

$$\sin\alpha \cong \frac{\lambda}{d} \tag{18}$$

which expresses [35, 39, 40, 65] the Rayleigh resolution criterion from classical optics. In (18) α denotes the "*angular resolution*", λ is the wavelength of light, and d represents the diameter of lens aperture. Note that criterion (18) was introduced in classical optics in 1879, i.e. by long time before the QM appeared. Later one relation (1) was introduced by taking in (18)$d \sim \Delta_{TE} \cdot q$ for coordinate uncertainty, respectively $\lambda = (\hbar/p)$ for momentum p (through wave-particle duality formula) and $p \cdot \sin\alpha \sim \Delta_{TE} \cdot p$ for momentum uncertainty.

3.9.2 D_{9b}: Classical "Gabor's uncertainty relation"

An example of an atypical analog of (2) can be found within the mathematical harmonic analysis in connection with a pair of random quantities regarded as Fourier conjugated variables (see [66, 67] and the Appendix C below). In non-quantum physics such an analogue is known [67] as "*Gabor's uncertainty relation*" which can be represented through the relation

$$\Delta t \cdot \Delta \nu \geqslant \frac{1}{4\pi}. \tag{19}$$

This last relation (19) shows the fact that for a classical signal, regarded as a wave packet (of acoustic or electromagnetic nature), the product of the "*uncertainties*" ("*irresolutions*") Δt and $\Delta\nu$ in the time and frequency domains cannot be smaller than a specific constant.

3.9.3 D_{9c}: A relation regarding thermodynamic observables

Another example of an atypical similar of UR (2) is given by the following classical formula

$$\Delta_W \mathbb{A} \cdot \Delta_W \mathbb{B} \geqslant \left|\langle \delta_W \mathbb{A} \cdot \delta_W \mathbb{B} \rangle_W\right| \tag{20}$$

showed as relation (D.3) in Appendix D of the present article. The elements (notations and physical significances) implied in (20) are those detailed in Appendix D. The respective elements are specific to the phenomenological theory, initiated by Einstein, about fluctuations of macroscopic thermodynamic observables (see [20, 68–72] and Appendix D below).

Note that, from the perspective of mathematics (more exactly of probability theory), the macroscopic formula (20) and UR (2) are analogue relations, both of them regard the fluctuations of the corresponding observables judged as random variables. Moreover they describe the intrinsic properties of considered systems (of macroscopic-thermodynamic respectively quantum nature) but not aspects of measurements performed on the respective systems. The corresponding measurements can be described through a distinct approaches modeled/depicted as transmission processes for stochastic data (see below Appendix E and Section 5 in present article).

As regards the formula (20), the following notifications should be done too. To a some extent the respective formula can be considered as being member to a family of so called "*thermodynamic UR*", discussed in a number of publications from the last century (see [78, 79] and references). Note that the alluded membership is true only in respect with the "*regular*" subset of respective family, derivable from the Einstein's phenomenological theory. But the mentioned family includes moreover a class of "*irregular*" relations. The most known such an "*irregular*" relation regards the conjugate variables energy U and temperature T of a thermodynamic system. It has [78] the form

$$\Delta U \cdot \Delta\left(\frac{1}{T}\right) \geqslant k_B \qquad (21)$$

where k_B denote the Boltzmann's constant.

It must be noted now the reality that fluctuation formula (20) and "*irregular*" relations like is (21) are completely dissimilar, first of all, due to the important distinction between reference frames of their definitions. The respective dissimilarity is pointed out by the following aspects. On the one hand, the quantities $\Delta_W \mathbb{A}$ and $\Delta_W \mathbb{B}$ from (20) are defined by referring to the same state of the considered system. On the other hand the quantities U and T which appear in (21) refer to different states of a system, namely states characterized by an energetic isolation respectively by a thermal contact. Due mainly to the above mentioned dissimilarity: "*a derivation of the uncertainty relation* (21) *analogous to that of the usual Heisenberg relations* (i.e. UR (2)) *is impossible*" [78].

Add here the fact that, within associate literature, it was reported a number of controversies about the aspects regarding the possible similarities between the "*thermodynamic UR*" (mainly from the same subset as (21)) and quantum UR (2) (see [78] and references). Among respective aspects can be quoted:

- compatibility of macroscopic observables,
- commutativity of thermodynamic variables and
- reconstruction of QM from hidden variables theories similarly with the rebuilding of thermodynamics through subjacent molecular considerations.

Note that the just mentioned aspects are not taken into account (as relevant elements) for our present survey on deficiencies of prevalent philosophy regarding UR and QMS.

3.9.4 D_{9d}: On the so called macroscopic operators

In the spirit of conventional precept ***BP*₅** the uniqueness of UR (2) consists in its strict specificity for micro-particles (of atomic size), without analogues in cases of macroscopic systems. But, as it is pointed out through relation (D.12) from Appendix D, in case of macroscopic thermodynamic system studied in quantum statistical physics one finds the formula

$$\Delta_\rho \mathbb{A} \cdot \Delta_\rho \mathbb{B} \geqslant \frac{1}{2}\left|\left\langle\left[\hat{\mathbb{A}}, \hat{\mathbb{B}}\right]\right\rangle_\rho\right|. \qquad (22)$$

This last formula is similar with the conventional UR (2) (more exactly, mathematically, with its primary versions (A.7) and (B.4)). Due to such a similarity, probably, some publications (e.g. [74] and references) have tried to regard (22) as a macroscopic UR. But the respective regard was found to be incompatible with the known UR–QMS philosophy, mainly with the precept ***BP*₄**.

The alluded incompatibility is pointed out by the following facts. On the one hand, in spirit of UR philosophy (precepts ***BP*₁–*BP*₄**), the quantities $\Delta_\rho \mathbb{A}$ and $\Delta_\rho \mathbb{B}$ from (22) should be considered as measuring uncertainties of macroscopic observables $\mathbb{A}$ and $\mathbb{B}$. Additionally when the operators $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$ and do not commute (i.e. $[\hat{\mathbb{A}}, \hat{\mathbb{B}}] \neq 0$), according to (22), the quantities $\Delta_\rho \mathbb{A}$ and $\Delta_\rho \mathbb{B}$ can be never reduced concomitantly to null values. Consequently, in terms of UR–QMS philosophy, for any situation, the non-commutable macroscopic observables $\mathbb{A}$ and $\mathbb{B}$ are allowed to be measurable simultaneously only with non-null and interconnected uncertainties. But, on the other hand, according to the classical physics any two macroscopic observables can be measured concurrently with unlimited accuracies and without any interrelated uncertainties.

For avoiding the above noted incompatibility some partisans of UR philosophy have suggested the following expedient. Abrogation of (22) by replacement of genuine macroscopic operators $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$ with another quasi-diagonal operators $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ (i.e. with operators whose representations in any base are quasi-diagonal matrices). Such substituting operators should to commute and so the right hand term in (22) to be (quasi) null (i.e. $\left|\langle[\hat{\mathbf{A}}, \hat{\mathbf{B}}]\rangle_\rho\right| \approx 0$). Through the mentioned substitution the inconvenient relation (22) could be changed with the more convenient formula

$$\Delta_\rho \mathbf{A} \cdot \Delta_\rho \mathbf{B} \geqslant \frac{1}{2}\left|\left\langle\left[\hat{\mathbf{A}}, \hat{\mathbf{B}}\right]\right\rangle_\rho\right| \approx 0. \qquad (23)$$

Then it seems to be possible that the substituted macroscopic uncertainties $\Delta_\rho \mathbf{A}$ and $\Delta_\rho \mathbf{B}$ to be reduced simultaneously to arbitrarily small (even zero) values. Apparently, such a possibility should to harmonize the interpretation of the relation (23) with the concepts of classical physics.

However, in fact, the above mentioned harmonization is not possible and the suggested expedient is useless. This, at least, due to the following reasons:

- Firstly, the relations (22) cannot be abrogated/substituted if the entire mathematical framework of quantum statistical physics is not abrogated/substituted too.
- Secondly, in common practice of studies of quantum statistical systems (e.g.such are the ones investigated in [80, 81]) are used the genuine operators $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$ but not the quasi-diagonal ones $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$.
- As a third reason, the following fact can be also noted. Even in certain situations when the original operators $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$ are quasi-diagonal in the sense of the mentioned expedient, the relation (23) does not turn into a form having a null term in the right hand side. Such a situation can be found [20] in case regarding a macroscopic paramagnetic system made of a huge number of independent 1/2-spins. In such a case as macroscopic operators appear the Cartesian components $\hat{\mathbb{M}}_\alpha$ ($\alpha = x, y, z$) of the system magnetization. Note that the

operators $\hat{\mathbb{M}}_\alpha$ are quasi-diagonal in the sense required by the aforesaid expedient/substitution. But, for all that, the respective operators do not commute because $[\hat{\mathbb{M}}_\alpha, \hat{\mathbb{M}}_\beta] = i\hbar\gamma \cdot \epsilon_{\alpha\beta\mu} \cdot \hat{\mathbb{M}}_\mu$ (γ = magneto-mechanical factor and $\epsilon_{\alpha\beta\mu}$ denotes the Levi-Civita tensor).

By taking into account the above pointed out deficiencies $\boldsymbol{D_9}$ (including $\boldsymbol{D_{9a}}$, $\boldsymbol{D_{9b}}$, $\boldsymbol{D_{9c}}$ and $\boldsymbol{D_{9d}}$) one may record the following conclusion. The relations (D.12)/(22) are relations regarding macroscopic areas of physics but not pieces which should be adapted to the requirements of prevalent philosophy about UR and QMS.

3.10 D_{10}: On the uniqueness of quantum measurements

Let us refer now to the uniqueness character of conventional relations (1) and (2) with regard to the measurements peculiarities at quantum level. The aforesaid character was largely debated in literature and it has generated the still open questions about the main characteristics (conceptual relevance and description procedures) of Quantum Measurements (QMS). By promoting all the assertions from percepts $\boldsymbol{BP_1}$–$\boldsymbol{BP_4}$ the UR–QMS philosophy tried to enforce the opinion that relations (1) and (2) are closely linked with the measuring particularities that are unique in quantum context, without any correspondence (analogy) in non-quantum domains of physics. The mentioned opinion, often promoted as a true dogma, dominates the mainstream of existing publications.

On the other hand, as we have argued above through the deficiencies $\boldsymbol{D_1}$–$\boldsymbol{D_9}$, the alluded opinion is completely unfounded because, genuinely, the respective relations are:

- either an old-fashioned (and removable) empirical convention (in case of (1)),
- or simple (non-magistral) theoretical formula (in case of (2)).

Within UR–QMS prevalent philosophy, as a widespread belief, the uniqueness peculiarities of QMS are motivated through the so called "observer effect". The respective effect is presented as a perturbing influence of observer (by experimental devices) on investigated systems and on measuring results. It is presumed to differentiate radically the QMS from classical measurements (of macroscopic physics). Such effects are absolutely unavoidable and affected by notable uncertainties in quantum contexts but entirely preventable and with negligible inaccuracies in classical situations.

The above mentioned belief is categorically disproved by the following observations. The "observer effect" appear not only in QMS but also in some classical measurements (e.g. [82] in electronics or in thermodynamics). Of course that in classical cases the measuring inaccuracies can be made negligible (by adequate improvements of experimental devices and/or procedures). It should be noted, that, in principle, quantum uncertainties can be also diminished (for example, with the super-resolution techniques discussed above in $\boldsymbol{D_1}$).

Then the idea of uniqueness quantum measuring character for conventional relations (1) and (2), promoted by UR philosophy through $\boldsymbol{BP_5}$, proves oneself as being a groundless fiction which should be disregarded. But such a disregard come to fortify the J. Bell's thinking [83, 84] that: "*the word 'measurement' should be avoided (or even … banned) altogether in quantum mechanics*". Some annotations about the respective thinking are given below in Section 5 where we will present briefly a non-conventional approach of QMS problems.

3.11 D_{11}: On the uniqueness of Planck's constant

Another aspect of quantum uniqueness invoked in precept $\boldsymbol{BP_5}$ regards the presence of Planck's constant $\hbar$ as a specific symbol in conventional quantum relations (1) and (2), comparatively with a total absence of some similar symbols in all classical (non-quantum) formulas. We shall examine the alluded aspect in regard with the relation (2). Then of prime importance is to notify the fact that, mathematically, quantum observables from the relation (2) have a stochastic (non-deterministic) character. But a completely similar character one finds in cases of macroscopic observables implied in formula (20) regarding fluctuations specific to macroscopic thermodynamic systems.

Both kinds of mentioned stochastic observables describe fluctuations (at quantum respectively macroscopic scale). The mentioned fluctuations are characterized quantitatively by the corresponding standard deviations such are $\Delta_\Psi A$ or $\Delta_W \mathbb{A}$. But, mathematically, the standard deviation indicates quantitatively the stochasticity (randomness) degree of an observable. This in the sense that the alluded deviation has a positive or null value as the corresponding observable is a random or, alternatively, a deterministic (non-stochastic) variable. Consequently the quantities $\Delta_\Psi A$ and $\Delta_W \mathbb{A}$ can be regarded as similar indicators of stochasticity for quantum respectively macroscopic observables.

In principle for macroscopic thermal fluctuations the standard deviations like is $\Delta_W \mathbb{A}$ can have various expressions (depending on system, state and observable). Apparently, it would seem that the respective expressions do not contain any common element. Nevertheless such an element can be found as being materialized by the Boltzmann's constant k_B (see relation (D.4) in Appendix D below and articles [71, 73]). So, for any macroscopic fluctuating observable A, the quantity $(\Delta_W \mathbb{A})^2$ (i.e. dispersion = square of the standard deviation) appears as a product of Boltzmann's constant k_B with factors which are independent of k_B.

This means that the quantity $(\Delta_W \mathbb{A})^2$, in its quality of quantitative indicator of thermal fluctuations, is directly proportional with k_B. Consequently $(\Delta_W \mathbb{A})^2$ has a non-null respectively null value as $k_B \neq 0$ or $k_B \to 0$ (Note that because k_B is a physical constant the limit $k_B \to 0$ means that the quantities directly proportional with k_B are negligible comparatively with other quantities of same dimensionality but independent of k_B). On the other hand, the standard deviation $\Delta_W \mathbb{A}$ is a particular indicator for macroscopic stochasticity revealed through thermal fluctuations.

Bringing together the above noted aspects it can be said that k_B has the qualities of an authentic generic indicator for thermal stochasticity which is specific for classical macroscopic fluctuating systems.

Now let us discuss about the quantum stochasticity whose particular indicators are the standard deviations $\Delta_\Psi A$. Based on the relations (13) one can say that in many situations the expressions for dispersions $(\Delta_\Psi A)^2$ consist in products of Planck constant $\hbar$ with factors which are independent of $\hbar$. Then, by analogy with the above discussed macroscopic situations, $\hbar$ places itself in the posture of generic indicator for quantum stochasticity.

The mentioned roles as generic indicators for k_B and $\hbar$ (in direct connections with the quantities $\Delta_W \mathbb{A}$ and $\Delta_\Psi A$) regard the one-fold (simple) stochasticity, of thermal and quantum nature respectively. But in physics is also known a twofold (double) stochasticity, of a combined thermal and quantum nature. Such a kind of stochasticity one finds in cases of macroscopic thermodynamic systems composed of statistical assemblies of quantum micro-particles. The alluded twofold stochasticity can be evaluated in a way through the dispersions $(\Delta_\rho \mathbb{A}_j)^2$ which estimate the level of fluctuations in the mentioned systems (see [20, 73, 76] and Appendix D below). As it is noted in relation (D.13) the dispersions $(\Delta_\rho \mathbb{A}_j)^2$ can be given through of products containing the function $\mathfrak{f}(k_B, \hbar) = \hbar \cdot \coth\left(\frac{\hbar\omega}{2k_B T}\right)$ and factors which are independent of both k_B and $\hbar$.

Then it results that k_B and $\hbar$ considered together turn out to be a couple of generic indicators for the twofold (double) stochaticity of thermal and quantum nature. Such a kind of stochaticity is significant or negligible in situations when $k_B \neq 0$ and $\hbar \neq 0$ respectively if $k_B \to 0$ and $\hbar \to 0$.

Now we can note the indubitable remark that Planck's constant $\hbar$ has an authentic classical analog represented by the Boltzmann's constant k_B, both $\hbar$ and k_B having relevant significances as generic indicators of stochaticity. But such an analogy contradicts directly the basic precept $\boldsymbol{BP_5}$ of CIUR doctrine and UR–QMS philosophy.

3.12 D_{12}: On the excessive ranking of UR

The ranking of UR to a position of principle, is widespread in the dominant literature, mainly through the authoritative and normative writings of many leading scientists. Surprisingly the respective ranking is argued merely in few occasions (e.g. in [10]) but only partially and not convincingly.

However, in [10], it was signaled the fact that "*over the years, some authors and foremost K. Popper, have contested this view*, of such a ... 'ranking' ". The mentioned contestation seems to have been motivated by the assertion: "*uncertainty relations cannot be granted the status of a principle on the grounds that they are derivable from the theory* ('QM'), *whereas one cannot obtain the theory from the uncertainty relations*". The aforesaid motivation was minimized and repudiated [10] through of the conventional (and prevalent) opinion that: "*there are many statements in physical theories which are called principles even though they are in fact derivable from other statements in the theory in question*". Note that in spite of the mentioned repudiation, it was added in [10] the noteworthy observation that "*Serious attempts to build up quantum theory as a full-fledged Theory of Principle on the basis of the uncertainty principle have never been carried out*".

As regards the above presented controversy our belief can be expressed as follows. The Popper's contestation of UR ranking (i.e., in fact, of the precept $\boldsymbol{BP_6}$) has a genuine character while the opposing conventional opinion is nothing but a questionable (and unfounded) attempt to preserve a predominant traditionalist doctrine (dogma).

Now, from another perspective, we wish to point out a new important aspect. On the one hand a true scientific conception attests indubitably the idea that: "*A principle is statement which is taken to be true at all times and all places where it is applicable*" [85]. On the other hand all previously proved deficiencies $\boldsymbol{D_1}$–$\boldsymbol{D_{11}}$ show that usual philosophy of UR is not valid in a wide class of situations where they should to be applied. Therefore such a philosophy cannot provide (generate) a principle (fundamental concept) applicable in an unquestionable manner for a large area of situations. That is why it turns out to be totally unacceptable (and useless) the idea to raise the entire UR philosophy to a rank of fundamental principle for QM.

Consequently, the precept $\boldsymbol{BP_6}$ shows oneself as being nothing but an unjustified thesis. At the same time, from a true scientific perspective, it is outside of acceptable usages to put in practice an idea such is [10]: "*we use the name "uncertainty principle" simply because it is the most common one in the literature*".

4 Which is really the true significance of UR?

Summing all the discussions incorporated within deficiencies $\boldsymbol{D_1}$–$\boldsymbol{D_{12}}$ one can notify the following evident remarks:

- There are profound deficiencies regarding all the basic elements and precepts of the conventional conceptions (CIUR doctrine and UR–QMS philosophy).
- In their essence the respective deficiencies are unavoidable and insurmountable within own framework of respective conceptions.
- Consequently the mentioned conceptions prove themselves as being undoubtedly in a failure situation which impose their abandonment.

The above argued abandonment of conventional conceptions points out very clearly the indubitable ending of the existing prevalent philosophy about UR. But a fair evaluation of such an ending requires an adequate epilogue regarding the future scientific status of the respective philosophy and of its constitutive and associate concepts.

The alluded epilogue demands firstly, detailed re-evaluations of the generative relations (1) and (2) from which have been expanded themselves the mentioned philosophy and concepts. The respective re-evaluations have to be done and argued by taking into account all the aspects noted previously within the texts of deficiencies $\boldsymbol{D}_1$–$\boldsymbol{D}_{12}$. Doing so one arrives to the following observations:

- Relation (1) is nothing but an old-fashioned (and removable) empirical convention. It persists as a piece of historical reminiscence, destitute of any wonderful status/significance for actual and future physics.
- Relation (2) proves to be only an ordinary QM formula, of well-defined (but not universal) validity. In such a posture it describes in a simple manner the connections between fluctuation characteristics of two quantum observables.
- In fact the relations (1) and (2) have not any crucial significance, for QM concretely and less so for physics in general.
- Relations (1) and (2) or their "adjustments" have not any connection with genuine descriptions of QMS.
- Particularly the respective relations do not depict in any way the so called "observer effect" (i.e. perturbing influence of "experimenter" on the investigated system).

5 Considerations on quantum measurements

Besides the main discussions about the meaning of early relations (1) and (2), the conventional UR philosophy generated also many collateral debates on Quantum Measurements (QMS) (see [1–12, 86–88] and references). The respective debates, still active in writings of many scientists, promoted an appreciable diversity of viewpoints about conceptual significance and practical importance of QMS. But in the same context, were recorded observations like is the following one

- "*Despite long efforts, no progress has been made … for … the understanding of quantum mechanics, in particular its measurement process and interpretation*" [89].

Nevertheless, beyond the mentioned debates, the respective subject of QMS involves also a matter of real interest for physics. That matter regards the natural interest in developing adequate theoretical description(s) for QMS, which should to be proved through viable arguments and which have to become of suitable utility for scientific and technical activities.

The above signaled situation have motivated interest for both conventional and non-conventional approaches of QMS problem. A modest non-conventional approach was put in work progressively in our investigations over many years (see [17–20, 47, 55, 90–94]). Here, as well as in all sections of present article, we try to gather, extend, systematize and improve the results of mentioned investigations in order to present argued viewpoints about the main aspects of QMS matter.

5.1 $\boldsymbol{D}_{13}$: The incorrect association of QMS with UR

As a first main aspect of the so much debated QMS problem is fact that it has a theoretical essence. Namely, it is focused around the idea of developing a general theoretical model for describing measurements on quantum systems. The respective model should have some similarity (a bit of reference) with the one centered on Schrödinger equation within QM.

From the perspective of the such supposed similarity most of publications promoted or accepted the opinion that QMS have a basic essentiality for QM in itself. During the years were recorded even assertions like the following one:

- '*the description of QMS is "probably the most important part of the theory (QM)*" ' [5].

But note that both the mentioned opinion and assertion are grounded on the belief that, mainly, the claimed essentiality/importance of QMS for QM is given by relations (1) and (2) in terms of precepts $\boldsymbol{BP}_1$–$\boldsymbol{BP}_6$.

On the other hand, it is easy to see that the respective belief is invalidated by the arguments from the entire collection of deficiencies $\boldsymbol{D}_1$–$\boldsymbol{D}_{12}$ notified by us above in Section 3.

Now, besides the aforesaid notifications, for starting our non-conventional approach of QMS subject, we take into account the following remarks of J. S. Bell:

- "*I agree with what you say about the uncertainty principle: it has to do with the uncertainty in predictions rather the accuracy of 'measurement'. **I think in fact that the word 'measurement' has been so abused in quantum mechanics that it would good to avoid it altogether***" (see [83] and Appendix I below).
- "*… The word ('measurement') has had such a damaging effect on the discussions that … it should be banned altogether in quantum mechanics*" [84].

A similar account we give also to the next remark:

- "*the procedures of measurement (comparison with standards) has a part which cannot be described inside the branch of physics where it is used*". [95]

The just noted remarks consolidate for us the following key view:

- The significance of UR is an intrinsic question of QM while the description of QMS constitutes an adjacent but distinct subject comparatively with QM in itself.

As another reference element for starting our approach we agree the following observation:

- "*it seems essential to the notion of measurement that it answers a question about the given situation existing before the measurement. Whether the measurement leaves the measured system unchanged or brings about a new and different state of that system is a second and independent question*" [96].

In sense of above observation for a measured physical system the "situation existing before the measurement" regards the intrinsic properties of that system. The characteristics of the respective properties play a role of input data (information) for measuring actions. On the other hand for the same system, the "answer (i.e. result) of measurement" is accumulated in "output data (information)" that are provided by measuring process. Correspondingly the whole measurement can be considered as a transmission process for information (stochastic data), while the measuring device appears as a communication channel (viewed as in [97]).

So the whole image of a measurement can be depicted through the scheme

$$\left| \begin{matrix} input \\ data \end{matrix} \right\rangle \Rightarrow \left[\begin{matrix} communication \\ channel \end{matrix} \right] \Rightarrow \left[\begin{matrix} output \\ data \end{matrix} \right]. \tag{24}$$

For giving concrete descriptions of the above scheme in cases of QMS (measurements on quantum systems) it should also to take into view the next remark

- "*To our best current knowledge the measurement process in quantum mechanics is non-deterministic*" [89].

In such a view the mentioned input and output data as well the description of a QMS have to be presented by means of some non-deterministic (stochastic or random) entities. For a measured quantum system the totality of input data can be considered as being comprised in its specific (intrinsic) wave function Ψ_{in}, with known stochastic/probabilistic own significance. As regards the same system the output data should be represented by some quantities having also stochastic features. Formally, such quantities can be considered as being incorporated in an output wave function Ψ_{out}. Then the measuring process appear as communication channel which transposes the wave function from a Ψ_{in} reading into a Ψ_{out} image. So it can be suggested that, in case of a QMS, the scheme (24) can be represented through the following generic pattern:

$$\left| \begin{matrix} \text{probabilistic} \\ \text{content of } \Psi_{\text{in}} \end{matrix} \right\rangle \Rightarrow \left[\widehat{SCC} \right] \Rightarrow \left[\begin{matrix} \text{probabilistic} \\ \text{content of } \Psi_{\text{out}} \end{matrix} \right] \tag{25}$$

where $\widehat{SCC}$, depicts the "***S****tochastic* ***C****ommunication* ***C****hannel*" regarded as an "*operator*" which describe the measuring process.

The above suggested pattern regarding QMS can be particularized for various concrete situations by using QM terminology. Two such particularization will be detailed below in the Subsections 5.2 and 5.4.

5.2 On an observable with discrete spectrum

Let us refer to the case of a QMS for a single quantum observable A endowed with a non-degenerate discrete spectrum of eigenvalues $\{a_j\}_{j=1}^{n}$. The respective observable is described by the operator $\hat{A}$ which satisfy the equations $\hat{A}\varphi_j = a_j \cdot \varphi_j$, where $\{\varphi_j\}_{j=1}^{n}$ signify the corresponding eigenfunctions.

If the set of eigenfunctions $\{\varphi_j\}_{j=1}^{n}$ is regarded as an orthonormal basis the wave functions Ψ_{in} and Ψ_{out} can be represented as follows

$$\begin{gathered} \Psi_{in} = \sum_{j=1}^{n} \alpha_j \varphi_j, \quad \sum_{j=1}^{n} \left|\alpha_j\right|^2 = 1, \\ \Psi_{out} = \sum_{k=1}^{n} \beta_k \varphi_k, \quad \sum_{k=1}^{n} |\beta_k|^2 = 1. \end{gathered} \tag{26}$$

Then the the pattern (25) appears as a transformation of the corresponding probabilities from *in*-readings $\{|\alpha_j|^2\}_{j=1}^{n}$ into *out*-images $\{|\beta_k|^2\}_{k=1}^{n}$. According to mathematics (probability and information theories) the mentioned transformation (i.e.the operator $\widehat{SCC}$) can be depicted by means of a doubly stochastic matrix M_{kj} $(k, j = 1, 2, \ldots, n)$, interpreted as in [98]. Such a depiction has the form

$$|\beta_k|^2 = \sum_{j=1}^{n} M_{kj} \cdot \left|\alpha_j\right|^2 \tag{27}$$

where the matrix M_{jk} satisfies the conditions

$$\sum_{k=1}^{n} M_{kj} = \sum_{j=1}^{n} M_{kj} = 1.$$

As above described a QMS appear as being ideal respectively non-ideal, accordingly as $M_{kj} = \delta_{kj}$ or $M_{kj} \neq \delta_{kj}$, where δ_{kj} denotes a Kronecker delta.

By using (26) and (27) for the η-expected values $\langle A \rangle_{\eta} = (\Psi_{\eta}, \hat{A}\Psi_{\eta})$, $(\eta = in, out)$, of observable A one obtains

$$\begin{gathered} \langle A \rangle_{in} = \sum_{j=1}^{n} a_j \cdot \left|\alpha_j\right|^2, \\ \langle A \rangle_{out} = \sum_{k=1}^{n} a_k \cdot |\beta_k|^2 = \sum_{k=1}^{n} \sum_{j=1}^{n} a_k \cdot M_{kj} \cdot \left|\alpha_j\right|^2. \end{gathered} \tag{28}$$

In terms of above notations the error for the expected value of A is:

$$\mathcal{E}\{\langle A \rangle\} = \langle A \rangle_{out} - \langle A \rangle_{in} = \sum_{k=1}^{n} \sum_{j=1}^{n} a_k \cdot \left(M_{kj} - \delta_{kj}\right) \cdot \left|\alpha_j\right|^2 \tag{29}$$

where δ_{jk} signifies a Kronecker delta.

Because, mathematically, the observable A is a random variable it is characterized also by the standard deviations

$\Delta_\eta A$ ($\eta = in, out$), defined as follows

$$\begin{aligned}(\Delta_{in}A)^2 &= \left\langle (A-\langle A\rangle_{in})^2\right\rangle_{in} \\ &= \sum_{j=1}^{n} a_j^2\cdot\left|\alpha_j\right|^2 - \left(\sum_{j=1}^{n} a_j\cdot\left|\alpha_j\right|^2\right)^2 \\ (\Delta_{out}A)^2 &= \left\langle (A-\langle A\rangle_{out})^2\right\rangle_{out} \\ &= \sum_{k=1}^{n}\sum_{j=1}^{n} a_k^2\cdot M_{kj}\cdot\left|\alpha_j\right|^2 \\ &\quad -\left(\sum_{k=1}^{n}\sum_{j=1}^{n} a_k\cdot M_{kj}\left|\alpha_j\right|^2\right)^2\end{aligned} \tag{30}$$

So for error $\mathcal{E}\{\Delta A\}$ of standard deviation regarding the observable A one finds

$$\begin{gathered}\mathcal{E}\{\Delta\, A\} = \Delta_{out}A - \Delta_{in}A \\ = \sqrt{\sum_{k=1}^{n}\sum_{j=1}^{n} a_k^2\cdot M_{kj}\cdot\left|\alpha_j\right|^2 - \left(\sum_{k=1}^{n}\sum_{j=1}^{n} a_k\cdot M_{kj}\left|\alpha_j\right|^2\right)^2} - \\ \sqrt{\sum_{j=1}^{n} a_j^2\cdot\left|\alpha_j\right|^2 - \left(\sum_{j=1}^{n} a_j\cdot\left|\alpha_j\right|^2\right)^2}\,.\end{gathered} \tag{31}$$

Now note the fact that, to some extent, the above presented model of a QMS description has general features. This because, excepting the conditions of being doubly stochastic, the measuring matrix M_{kj} can consists of arbitrary components. The mentioned generality/arbitrariness should be reduced when one refers to the relatively accurate measurements. Such a reduction can be modeled if the measuring matrix elements M_{kj} are taken of the forms

$$\begin{gathered}M_{kj} = \delta_{kj} + \tau_{kj}, \\ \left|\tau_{kj}\right| << 1, \qquad \sum_{k=1}^{n}\tau_{kj} = \sum_{j=1}^{n}\tau_{kj} = 0,\end{gathered} \tag{32}$$

where δ_{kj} signifies the a Kronecker delta and τ_{kj} are real and dimensionless quantities of (very) small values.

When the matrix elements M_{jk} are approximated as in (32) the errors $\mathcal{E}\{\langle A\rangle\}$ and $\mathcal{E}\{\Delta\, A\}$ from (29) and (31) can be estimated through a direct calculation, respectively by means of the first order term in Taylor series. Then one finds

$$\begin{gathered}\mathcal{E}\{\langle A\rangle\} = \sum_{k=1}^{n}\sum_{j=1}^{n} a_k\cdot\tau_{kj}\cdot\left|\alpha_j\right|^2, \\ \mathcal{E}\{\Delta A\} \approx \sum_{k=1}^{n}\sum_{j=1}^{n}\left[\frac{\partial\mathcal{E}(\tau_{kj})}{\partial\tau_{kj}}\right]_{\tau_{kj}=0}\cdot\tau_{kj},\end{gathered} \tag{33}$$

where $\mathcal{E}\left(\tau_{kj}\right)$ signifies the standard-deviation error $\mathcal{E}\{\Delta\, A\}$ from (31) in which one uses the approximations (32).

Relations (33) show that within mentioned approximations the parameters τ_{jk} appear as significant indexes regarding the measuring accuracies. So the discussed measurement can be regarded as ideal when $\tau_{kj} = 0$ for all k and j, respectively as non-ideal when $\tau_{kj} \neq 0$ at least for some values of k or j.

5.3 $\boldsymbol{D_{14}}$: On the measuring scenarios with a unique sampling

As it was pointed out in Subsection 5.1, a QMS is essentially a non-deterministic process. Due to the mentioned essentiality, the "result" of such a process must be represented in terms of some stochastic (probabilistic) output data. But, surprisingly, in conventional publications [99–106] a QMS is regarded as a scenario (i.e. an imagined sequence of possible events) conceived as a single sampling (i.e. as a unique-deterministic selection from a set of random data). So regarded, a QMS gives as its result (outcome) a single value in which falls (collapses) the whole physical content of the measured observable. The referred falling scenarios are illustrated by two widely debated themes regarding the Wave Function Collapse (WFC) [99–103] respectively the Schrödinger's Cat Thought Experiment (SCTE) [104–106]. Historically, both the respective themes have occurred in a direct connection with the establishing of basic precepts ***BP*₁–*BP*₆** of CIUR doctrine and UR–QMS philosophy. Therefore, by taking into account the deficiencies of precepts ***BP*₁–*BP*₆**, revealed above in Section 3, it is here the place to investigate also the possible deficiencies of the aforesaid scenarios.

Let us begin the announced investigation by referring to the WFC-measuring-scenario. The respective scenarios germinated from the hypothesis that, due to unavoidable measuring perturbations, all QMS cause specific collapses (falls, jumps) in states of the measured quantum systems. It can be presented succinctly in usual terms of QM as follows.

Consider a measuring investigation focused on the system and observable A discussed in the previous Subsection 5.2. For the respective system in WFC-scenario the "situation existing before measurement" is inscribed in its intrinsic wave function Ψ_{in}. The probabilistic content of Ψ_{in} play the role of input data (information) for investigation actions. But, attention, within the WFC-scenario, the measuring actions are imagined as providing as result an ***u***nique ***d***etertministic ***o***utcome (***udo***) namely a_k.

Note that a_k is one of the eigenvalues $\{a_j\}_{j=1}^{n}$ from the spectrum of A. The eigenvalues $\{a_j\}_{j=1}^{n}$ are defined through the relations $\hat{A}\varphi_j = a_j\cdot\varphi_j$ ($j = 1, 2, \ldots, n$), where $\{\varphi_j\}_{j=1}^{n}$ denote the eigenfunctions of operator $\hat{A}$ associated to the observable A. Then, in terms detailed previously in Subsection 5.2, the whole WFC-scenario can be illustrated through the following two schemes

$$\left|\{a_j\}_{j=1}^{n}\cup\left\{\left|\alpha_j\right|^2\right\}_{j=1}^{n}\right\rangle \Rightarrow \left[\widehat{\boldsymbol{udo}}\right] \Rightarrow a_k, \tag{34}$$

$$\left|\Psi_{in} = \sum_{j=1}^{n} \alpha_j \cdot \varphi_j\right\rangle \Rightarrow \left[\widehat{\mathfrak{udo}}\right] \Rightarrow \varphi_k, \tag{35}$$

where $\widehat{\boldsymbol{udo}}$ symbolize an operator which describe the mesuring actions in WFC-scenario.

On the one hand, firstly, the schema (34) regards the measurement of observable A. It show a falling of the respective observable from a whole spectrum of values $\{a_j\}_{j=1}^{n}$, having probabilities $\{\left|\alpha_j\right|^2\}_{j=1}^{n}$ in measured state, to a unique value a_k as result of the scenario. Secondly, on the other hand, the schema (35) refers to the evolution of the considered system from a state "existing before the measurement" (at the beginning of scenario) in an "after measurement" state (in the end of scenario).

Specify here the fact that conventional publications (see [99–103] and references) regard relation (35) as being the essential symbol of WFC. That is why the mentioned publications tried to done analytical representations of the respective relation considered as image of a dynamical physical process. For such representations were promoted various inventions, e.g. nonlinear extensions of Schrödinger equation or even appeals to new kinds of fundamental physical constants.

The above mentioned WFC-scenario regarding QMS can be admonished through the following remarks.

Firstly note that quantum observables are stochastic variables. Consequently a true measurement of such an observable should be regarded as being provided not by an ***udo*** (***u**nique **d**eterministic **o**utcome*) but by an adequate probabilistic set of such outcomes. The data given by the respective set are expected to provide relevant (and as complete as possible) information about the considered observables.

Secondly, the idea of describing QMS through an analytical representation of the WFC schema (35) proves oneself as being an extravagance without solid arguments or credible hypotheses. Some main aspects of the respective extravagance can be revealed by taking into account the stochastic similitude between quantum and thermal (macroscopic) random observables. Such a reveal we point out here as follows.

Let us refer to a macroscopic thermodynamic system described in terms of phenomenological theory of fluctuations (see below the Appendix D). For simplicity the system will be considered to be characterized by a single macroscopic thermodynamic observable $\mathbb{A}$. Mathematically the macroscopic fluctuations of $\mathbb{A}$ are accounted by a real random variable $\mathcal{A}$ and described by the probability density $W = W(\mathcal{A})$. Through the before specified terms can be pointed out the analogy between measuring acts regarding the stochastic observables of quantum and macroscopic nature. An ***udo***, specific to WFC-scenario, for a quantum observable was discussed succinctly above in connection with the relations (34) and(35). A completely similar ***udo*** regarding a macroscopic observable $\mathbb{A}$ can be depicted as follows. By means of an ***udo*** for the variable $\mathcal{A}$ one obtains a unique value say $\mathcal{A}_0$. Then for $\mathbb{A}$ the respective ***udo*** can be illustrated through the following relations

$$|\mathcal{A} \in (-\infty, +\infty)\rangle \Rightarrow \left[\widehat{\boldsymbol{udo}}\right] \Rightarrow \mathcal{A}_0, \tag{36}$$

$$|W(\mathcal{A})\rangle \Rightarrow \left[\widehat{\boldsymbol{udo}}\right] \Rightarrow \delta(\mathcal{A} - \mathcal{A}_0), \tag{37}$$

where $\delta(X)$ denotes the Dirac's δ-function of X.

In principle, the aspects of quantum and macroscopic observables, depicted by (34) and (35) respectively (36) and (37) are completely similar. Therefore the discussions regarding the two kinds of ***udo*** should be similarly too. But in the macroscopic case the relation (37) is not considered at all as illustrating a dynamic process. Moreover within the corresponding macroscopic studies there is no interest for giving an analytical representation (through some evolution equations) regarding a scenario of type (37). This even if for the investigation of macroscopic observables one can use in principle a subjacent description given by classical statistical mechanics. Then, by virtue of above noted similarity, it can be said that the quantum scenario (35) should be not considered as a dynamic process. Consequently the QM studies have to be not concerned about the analytical representation (by some evolution equations) of an ***udo*** as the one illustrated by (35). Such regards about the scenario (35) are required, with all the more, as QM is not complemented (until today) by any subjacent theory of sub-quantum essence. Furthermore, for a true physical approach, the result of the respective ***udo*** must be gathered together with the answers of a significant statistical group of many other akin ***udo***. The respective answers should allow to find adequate probabilistic estimators of the investigated quantum observable.

Regarding the problem of QMS description, in the category of falling scenarios, along with the WFC idea one finds also the famous problem of SCTE (Schrödinger's Cat Thought Experiment). The respective problem, known also as Schrödinger's cat paradox, has retained the attention of many debates over the decades (see [104–106] and references). The essential element in SCTE is represented by a single decay of an individual radioactive atom (which, through some macroscopic machinery, kills an initially living cat). But the individual lifetime of a single decaying atom is a stochastic (random) variable. That is why the mentioned killing decay is in fact a twin analogue of the above mentioned ***udo*** taken into account by the WFC-scenario. So, the above considerations reveal the notifiable fact that, for a true evaluation of a stochastic observable (such is the mentioned decay lifetime), is worthlessly to operate with an ***udo*** which gives an unique result of measurement. Accordingly, the SCTE problem appears as a twin analogue of the IWFC-scenario, i.e. as a fiction (figment) without any real scientific value.

The aforesaid fictional character of the SCTE can be pointed out once more by observation [93, 94] that it is possible to imagine a macroscopic thought-experiment completely analogous with the SCTE. Within the respective macroscopic analogue, a cousin of Schrödinger's cat can be killed through launching a single macroscopic ballistic projectile. More specifically, the killing machinery is activated by an uncontrollable (unobservable) sensor located within the "circular error probable" (CEP) [109] of a ballistic projectile trajectory. The hitting point of the projectile is expected to arrive within CEP with the probability 50%. That is why the murderous action of a single launched projectile is just as much unpredictable as that of the unique radioactive atom within original SCTE. Therefore, the mentioned macroscopic analogy makes clear once more the fictional character of the SCTE.

According to the above-noted remarks, it should be regarded as worthless statements some assertions such as: "*The Schrödinger's cat thought experiment remains a defining touchstone for modern interpretations of quantum mechanics*" [106]. Note that such or similar assertions can be found in many popular publications or in the texts disseminated via the Internet (e.g. [110]).

Therefore SCTE problem as well as its similar WFC idea, discussed previously, prove themselves to be not real scientific topics but rather useless exercises (fictive scenarios), without any conceptual or practical significance.

5.4 About observables with continuous spectra

As it was noted in the beginning of this Section 5, for physics, development of suitable models for QMS description present a natural necessity. Above, in Subsection 5.2 of this article, it is detailed such a model regarding the measurement of an observable endowed with a discrete non-degenerate spectra. Here below we try to propose a measuring model with similar purpose (QMS description) but regarding observables having continuous spectra of values.

As in case with discrete spectrum for here regarded measuring situation we adopt the same generic pattern depicted in (25). The probabilistic content of wave functions Ψ_{in} and Ψ_{out} incorporate information (data) about the intrinsic state of the measured system respectively concerning the results provided by measurement. We will restrict our considerations to the measurements of orbital characteristics for a quantum spin-less micro-particle, supposed in a unidirectional motion along the x-axis. Note that the announced considerations can be easily extended for measurements regarding systems with spatial orbital motions. Then the wave functions Ψ_η ($\eta = in, out$) will be taken of the form $\Psi_\eta = \Psi_\eta(x)$ (note that here we omit to specify the time t as visible variable because the considered state of system refers to a given ante-measurement instant).

Note now the fact that according QM rules the wave functions Ψ_η have only significance of probability amplitudes but not a direct probability meaning. Therefore, in the case of interest here, the picture (25) of QMS should be detailed not in terms of wave functions Ψ_η, but by means of some entities with direct probabilistic meanings. This especially because the real measuring devices report the occurrence of some random values for investigated observables. In usual terms of QM entities with direct probabilistic significance are carriers of stochasticity: probability densities ρ_η and probability currents j_η ($\eta = in, out$). Let us write the wave functions Ψ_η as $\Psi_\eta(x) = \left|\Psi_\eta(x)\right| \cdot \exp\{i\,\Phi_\eta(x)\}$. Then, for a micro-particle with mass m considered as measured system, the alluded ρ_η and j_η are given by relations:

$$\begin{aligned} \rho_\eta &= \rho_\eta(x) = \left|\Psi_\eta(x)\right|^2, \\ j_\eta &= j_\eta(x) = \frac{\hbar}{m}\left|\Psi_\eta(x)\right|^2 \cdot \nabla_x \Phi_\eta(x), \end{aligned} \tag{38}$$

where $\nabla_x = \frac{\partial}{\partial x}$.

Now it must to specify that ρ_η and j_η refer to the positional and the motional kinds of probabilities respectively. Experimentally the two kinds can be regarded as measurable by distinct devices and procedures. The situation is similar with that of electricity studies where the aspects regarding position and mobility of electrical charges are evaluated through completely different devices and procedures. Due to the aforesaid specifications it results that in fact the generic pattern depicted in (25) has to be amended as follows

$$|\rho_{in}(x) \cup j_{in}(x)\rangle \Rightarrow \left[\widehat{SCC}\right] \Rightarrow \left[\rho_{out}(x) \cup j_{out}(x)\right]. \tag{39}$$

Mathematical considerations about the relations (25) and (E.1), (early referred also in [107]) can be applied by similarity for the pattern (39). So the respective pattern (i.e. the operator $\widehat{SCC}$)can be represented through the next two transformations:

$$\begin{aligned} \rho_{out}(x) &= \int_{-\infty}^{+\infty} \Gamma(x, x') \cdot \rho_{in}(x') \cdot dx', \\ j_{out}(x) &= \int_{-\infty}^{+\infty} \Lambda(x, x') \cdot j_{in}(x') \cdot dx'. \end{aligned} \tag{40}$$

Here $\Gamma(x, x')$ and $\Lambda(x, x')$ represent the corresponding doubly stochastic kernels (in sense defined in [108]). This means that the kernels $\Re = \{\Gamma,\ \Lambda\}$ satisfy the following relations $\int_{-\infty}^{+\infty} \Re(x, x')\, dx = \int_{-\infty}^{+\infty} \Re(x, x')\, dx' = 1$. The mentioned kernels incorporate some extra-QM elements regarding the characteristics of measuring devices and procedures. Such elements do not belong to the usual QM framework which refers to the intrinsic (own) characteristics of the measured micro-particle (system).

Through the above considerations can be evaluated the effects induced by QMS. The respective effects regards the probabilistic estimators for orbital observables A_j of considered quantum system. Such observables are described by the operators $\hat{A}_j$ ($j = 1, 2, \ldots, n$). As in case of classical measuring model (see the Appendix E), without any loss of generality, here one can suppose that the quantum observables have identical spectra of values in both *in*-and *out*-situations. In terms of QM the mentioned supposition means that the operators $\hat{A}_j$ have the same mathematical expressions in both *in*- and *out*-readings, i.e. that the respective expressions remain invariant under the transformations which describe QMS. In the here discussed case of a system with rectilinear orbital motion the mentioned expressions depend on x and ∇_x.

So one can say that in the situations associated with the wave functions $\Psi_\eta = \Psi_\eta(x)$ ($\eta = in, out$) the mentioned quantum observables A_j, can characterized by the following lower order estimators (or numerical parameters): mean values $\langle A_j\rangle_\eta$, correlations $C_\eta(A_j, A_k)$ and standard deviations $\Delta_\eta A_j$. We use the common notation (f, g) for scalar product of functions f and g, i.e. $(f, g) = \int\limits_{-\infty}^{+\infty} f^*(x) \cdot g(x) \cdot dx$. Then the mentioned estimators are defined by the relations

$$\begin{aligned} &\left\langle A_j\right\rangle_\eta = \left(\Psi_\eta, \hat{A}_j\Psi_\eta\right), \qquad \delta_\eta\hat{A}_j = \hat{A}_j - \left\langle A_j\right\rangle_\eta, \\ &C_\eta\left(A_j, A_k\right) = \left(\delta_\eta\hat{A}_j\Psi_\eta, \delta_\eta\hat{A}_k\Psi_\eta\right), \\ &\Delta_\eta A_j = \sqrt{C_\eta\left(A_j, A_j\right)}. \end{aligned} \tag{41}$$

Note here the fact that, on the one hand, the *in*-version of discussions the estimators (41) are calculated by means of the wave function Ψ_{in}. The respective function is supposed as being known from the considerations about the intrinsic properties of the investigated system (e.g. by solving the corresponding Schrödinger equation).

On the other hand, apparently, the evaluation of estimators (41) in η= *out*-version requires to operate with the wave function Ψ_{out}. But the respective appearance can be surpassed [20] through operations which use the probability density ρ_{out} and current j_{out}. So if an operator $\hat{A}_j$ does not depend on $\nabla_x = \frac{\partial}{\partial x}$, i.e. $\hat{A}_j = \hat{A}_j(x)$, in evaluating the scalar products from (41) can be used the evident equality $\Psi^*_{out}\hat{A}_j\Psi_{out} = \hat{A}_j \cdot \rho_{out}$. Additionally, when $\hat{A}_j$ depends on $\nabla_x = \frac{\partial}{\partial x}$, i.e. $\hat{A}_j = \hat{A}_j(\nabla_x)$, in the same products the expressions of the type $\hat{A}_j(\nabla_x)\Psi_{out}(x)$ can be converted in terms of $\rho_{out}(x)$ and $j_{out}(x)$. Namely from (38) one finds directly:

$$\begin{aligned} \nabla_x \left|\Psi_{out}(x)\right| &= \nabla_x \sqrt{\rho_{out}(x)}, \\ \nabla_x \Phi_{out}(x) &= \frac{m}{\hbar}\frac{j_{out}(x)}{\rho_{out}(x)}. \end{aligned} \tag{42}$$

By a single or repeated application of these formulas, any expression of type $\hat{A}_j(\nabla_x)\Psi_{out}(x)$ can be transcribed in terms of ρ_{out} and j_{out}.

The aforesaid discussion should be supplemented by specifying some indicators able to characterize the errors (uncertainties) of considered QMS. For the above quoted observables A_j such indicators are the following ones:

$$\left.\begin{aligned} &\mathcal{E}\left\{\left\langle A_j\right\rangle\right\} = \left\langle A_j\right\rangle_{out} - \left\langle A_j\right\rangle_{in} \\ &\mathcal{E}\left\{C\left(A_j, A_k\right)\right\} = C_{out}\left(A_j, A_k\right) - C_{in}\left(A_j, A_k\right) \\ &\mathcal{E}\left\{\Delta A_j\right\} = \Delta_{out}A_j - \Delta_{in}A_j \end{aligned}\right\} \tag{43}$$

The above presented model regarding the description of QMS for observables with continuous spectra is illustrated on a simple example in the Appendix G below.

6 Some concluding remarks

The present paper was motivated by the existence of many unclearnesses (unfinished controversies and unelucidated questions) about the prevalent UR–QMS philosophy. It was built as a survey on deficiencies of respective philosophy. So were re-evaluated the main ideas claimed within the mentioned philosophy. The basic results of the respective re-evaluations can be summarized through the following ***C***oncluding ***R***emarks (***CR***):

- ***CR*$_1$**: Firstly, through multiple arguments, we have proved the observation that the UR (1) and (2) have not any essential significance for physics. Namely the respective UR are revealed as being either old-fashioned, short-lived (and removable) conventions (in empirical, thought-experimental justification) or simple (and limited) mathematical formulas (in theoretical vision). But such an observation comes to advocate and consolidate the Dirac's intuitive prediction [23]: "*I think one can make a safe guess that uncertainty relations in their present form will not survive in the physics of future*". Note that the respective prediction was founded not on some considerations about the UR essence but on an intuition about the future role in physics of Planck's constant $\hbar$. Dirac predicted that $\hbar$ will become a derived (secondary) quantity while c and e will remain as fundamental constants (c = speed of light and e = elementary electric charge). ■

- ***CR*$_2$**: A significant idea that emerges from previous discussions is the one that neither UR (1) and (2) nor various "generalizations" of them, have not any connection with genuine descriptions of Quantum measurements (QMS). All the respective descriptions should be considered as a distinct (and additional) subject which must be investigated separately but somewhat in association with QM. Examples of such description are presented briefly, in Subsection, 5.2 and 5.4, for observables having discrete respectively continuous spectra. ■

- ***CR*$_3$**: Note that, in all of their aspects, the discussions from Subsection 5.2 and 5.4 have a theoretical essence. This means that, the entities like wave function Ψ_{in} as well as the measuring indicators M_{jk}, $\Gamma(x, x')$ and $\Lambda(x, x')$, are nothing but abstract concepts which enable elaboration of theoretical models regarding the descriptions of QMS. On the one hand Ψ_{in} refers to the intrinsic data about the studied system. It is evaluated by means of some known theoretic procedures (e.g. by means of the corresponding Schrödinger equation). On the other hand the indicators M_{jk}, $\Gamma(x, x')$ and $\Lambda(x, x')$ are introduced as theoretical entities for modeling the characteristics of the considered measuring devices/processes. ■

• $\boldsymbol{CR_4}$: Correlated with the previous $\boldsymbol{CR_2}$ and $\boldsymbol{CR_3}$ it must be specified that, in relation with QMS, the inventions of Wave Function Collapse (WFC) and Schrödinger's Cat Thought Experiment (SCTE) are nothing but unnatural falling scenarios. Consequently, as we have argued above in Subsections 5.3, both idea of WFC and SCTE problem prove themselves as being not real scientific subjects but rather unnecessary figments.■

• $\boldsymbol{CR_5}$: It is interesting to note here the fact that the history of quantum mechanics was abounded by an impressive number of publications related to UR–QMS philosophy. So, for the years between 1935 and 1978, as regards EPR (Einstein-Podolsky-Rosen) paradox, associated [112] with the situation of non-commuting observables, some authors [113] noted that "$\geqslant 10^6$ *papers have been written*" — i.e. $\geqslant$ 63 papers per day (!?). Also the same publishing abundance about QM matters (including UR–QMS philosophy) motivates remarks such are the following ones: "*A theory whose formalism can be written down on a napkin whilst attempts to interpret it fill entire libraries. A theory that has seen astonishing experimental conformation yet leaves us increasingly perplexed the more we think about it. How can we know so well how to apply this theory but disagree so vehemently about what it is telling us?*" [114]. Probably that, in some future, the alluded abundance will be investigated from historic and sociological perspectives.

• $\boldsymbol{CR_6}$: Over the years original UR (1) and (2) were supplemented with many kinds of "generalizations" (see [115–120] and references). Until today, the respective "generalizations" appear as being de facto only extrapolation mathematical "constructs" (often of impressive inventiveness). As a rule, they are not pointed out as having significance for some concrete physical questions (of conceptual or experimental relevance). But the existence of such significance is absolutely necessary in order to associate the mentioned "generalizations" with matters of certain importance for physics. In the light of the discussions from the present paper one can say that the sole physical significance of some from the referred "generalizations" seems to be their meaning as quantitative indicators of fluctuations (i.e. of stochasticity). But from a practical perspective among the respective indicators of practical usage are only the ones of relative lower order. Therefore, for tangible interests of physics, all the discussed "generalizations" seem to be rather excessive pieces. They remain only as interesting mathematical "constructs", which ignore the desideratum: "*Entities are not to be multiplied beyond necessity*". ■

• $\boldsymbol{CR_7}$: In discussions and revaluations proposed in this article, we have referred only to the prevalent philosophy of UR and QMS regarding primarily the foundations and interpretation of QM. But, as it is known, the mentioned philosophy has been extrapolated in other "extra muros" domains, outside of QM. As aforesaid domains can be quoted the following ones: mathematical computations, biology and medical sciences, economy and finance, human behavior, social sciences and even politics. A relevant bibliography regarding the mentioned extramural extrapolations can be accessed easy via internet from Google sites. Note that our above reevaluations of UR–QMS philosophy do not contain analyzes referring to the mentioned extrapolations. Such analyzes remain as task for scientists working in the respective domains. Here we want to point out only one noticeable aspect that differentiates the extramural UR from the primary ones. On the one hand, according to their origin, the primary UR from QM are strongly associated with a cardinal marker represented by the Planck constant $\hbar$. On the other hand, as far as we know, for extramural extrapolations of UR, the existence of similar markers, represented by cardinal indicators of the corresponding scientific domains, were not reported. ■

• $\boldsymbol{CR_8}$: In their essence, the above argued revaluations of UR and QMS, do not disturb in any way the basic framework of usual QM. This means that QM keeps its known specific elements: concepts (wave functions, operators) with their significances (of stochastic essence), principles and theoretical models (Schrödinger equation), computing rules (exact or approximate) and investigate systems (atoms, molecules, mesoscopic structures). Note here the observation that, for nowadays existing quantum debates, the above revaluations of UR–QMS, offer a few arguments for lucrative parsimony in approaches of matters. The unlucrative aspects of those debates have to be reconsidered too, probably in more or less speculative visions. We recall here that the basic framework of QM can be deduced [121] from direct physical considerations, without appeals to ambiguous discussions about UR or QMS. The alluded considerations start from real physical facts (particle-wave duality of atomic size systems). Subsequently they use the continuity equations for genuine probability density and current. After that one obtains the whole framework of QM (i.e. the Schrödinger equation, expressions of operators as descriptors of quantum observables and all the practical rules of QM regarded as a theoretical model for the corresponding investigated systems).

In the mentioned perspective, we dare to believe that, to some extent, the revaluations of UR and QMS promoted by us can give modest support for genuine reconsiderations regarding the interpretation and foundations of QM.■

Appendices

A: A brief compendium of some QM elements

Here we remind briefly some significant elements, selected from the usual theoretical framework [5, 29, 30] of Quantum Mechanics (QM). In this appendix we use Traditional Notations (TN), taken over from mathematical algebra developed long before QM appeared. Few specifications about the more recent Dirac's braket formalism are given in Appendix B.

So, in terms of TN, we consider a QM micro-particle whose state (of orbital nature) is described by the wave function Ψ. Two observables A_j $(j = 1, 2)$ of the respective particle will be described by the operators $\hat{A}_j$. The notation (f, g) will be used for the scalar (inner) product of the functions f and g. Correspondingly, the quantities $\langle \hat{A}_j \rangle_\Psi = (\Psi, \hat{A}_j \Psi)$ and $\delta_\Psi \hat{A}_j = \hat{A} - \langle \hat{A}_j \rangle_\Psi$ will depict the mean (expected) value respectively the deviation-operator of the observable A_j regarded as a random variable. Then, by denoting two observables with $A_1 = A$ and $A_2 = B$, one can be written the following formula:

$$\begin{aligned}\left(\delta_\Psi \hat{A}\Psi, \delta_\Psi \hat{A}\Psi\right) \cdot \left(\delta_\Psi \hat{B}\Psi, \delta_\Psi \hat{B}\Psi\right) &\geqslant \\ \geqslant \; \left|\left(\delta_\Psi \hat{A}\Psi, \delta_\Psi \hat{B}\Psi\right)\right|^2 &\end{aligned} \tag{A.1}$$

which is nothing but a relation of Cauchy-Schwarz type from mathematics.

For an observable A_j considered as a random variable, in a mathematical sense, the quantity $\Delta_\Psi A_j = (\delta_\Psi \hat{A}_j \Psi, \delta_\Psi \hat{A}_j \Psi)^{1/2}$ signifies its standard deviation. From (A.1) it results directly that the standard deviations $\Delta_\Psi A$ and $\Delta_\Psi B$ of the mentioned observables satisfy the formula

$$\Delta_\Psi A \cdot \Delta_\Psi B \geqslant \left|\left(\delta_\Psi \hat{A}\Psi, \delta_\Psi \hat{B}\Psi\right)\right|. \tag{A.2}$$

This last formula, with quantities $\Delta_\Psi A$ and $\Delta_\Psi B$ regarded together, play an influential role in QM debates within UR and QMS philosophy. That is why the relation (A.2) can be called *Cauchy-Schwarz Quantum Formula* (CSQF). Note that formulas (A.1) and (A.2) are always valid, i.e. for all observables, particles and states. Therefore they must be considered as primary QM formulas.

For the discussions regarding the UR–QMS philosophy it is helpful to present the particular versions of formula (A.1) in the cases when the operators $\hat{A} = \hat{A}_1$ and $\hat{B} = \hat{A}_2$ satisfy the conditions

$$\text{iff}: \quad \left(\hat{A}_j\Psi, \hat{A}_k\Psi\right) = \left(\Psi, \hat{A}_j\hat{A}_k\Psi\right), \quad (j, k = 1, 2) \tag{A.3}$$

(where iff ≡ if and only if). In the alluded cases it is true the next formula

$$\begin{aligned}\left(\delta_\Psi \hat{A}\,\Psi, \delta_\Psi \hat{B}\,\Psi\right) &= \tfrac{1}{2}\left(\Psi, \left\{\delta_\Psi \hat{A}\,, \delta_\Psi \hat{B}\right\}\Psi\right) \\ &\quad -\tfrac{i}{2}\left(\Psi, i\left[\hat{A}, \hat{B}\right]\Psi\right).\end{aligned} \tag{A.4}$$

Here $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$ and $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ signify the anticommutator respectively commutator of the operators $\hat{A}$ and $\hat{B}$. Now note the fact that the two terms from the right hand side of (A.4) are purely real and strictly imaginary quantities respectively. Therefore in the mentioned cases from (A.2) follows directly the enlarged inequality

$$\begin{aligned}(\Delta_\Psi A)^2 \cdot (\Delta_\Psi B)^2 \geqslant\; & \tfrac{1}{4}\left|\left\langle\left\{\delta_\Psi \hat{A}, \delta_\Psi \hat{B}\right\}\right\rangle_\Psi\right|^2 \\ & +\tfrac{1}{4}\left|\left\langle\left[\hat{A}, \hat{B}\right]\right\rangle_\Psi\right|^2.\end{aligned} \tag{A.5}$$

Sometimes this relation is referred to as the Schrödinger inequality. It imply subsequently the next two truncated inequalities

$$\Delta_\Psi A \cdot \Delta_\Psi B \geqslant \frac{1}{2}\left|\left\langle\left\{\delta_\Psi \hat{A}, \delta_\Psi \hat{B}\right\}\right\rangle_\Psi\right|, \tag{A.6}$$

$$\Delta_\Psi A \cdot \Delta_\Psi B \geqslant \frac{1}{2}\left|\left\langle\left[\hat{A}, \hat{B}\right]\right\rangle_\Psi\right|. \tag{A.7}$$

One observes that (A.7) is nothing more than the conventional Robertson-Schrödinger relation (2), commonly quoted in the literature of CIUR doctrine and UR–QMS philosophy. Note that in the respective literature besides the relation (2)/(A.7) sometimes the formula (A.5) is also mentioned. But, as a fact, the respective mention is not accompanied with the important specification that formula (A.5) is valid iff (if and only if) the condition (A.3) is fulfilled.

In the end of this appendix we note the cases of more than two observables, i.e. for a set A_j $(j = 1, 2, \ldots, n;\ n \geqslant 3)$, when the quantities $\alpha_{jk} = (\delta_\Psi \hat{A}_j \Psi, \delta_\Psi \hat{A}_k \Psi)$ constitute the components of a positive semi definite matrix. In such cases, similarly with (A.1), are true the formulas

$$\det\left[\left(\delta_\Psi \hat{A}_j\Psi, \delta_\Psi \hat{A}_k\Psi\right)\right] \geqslant 0; \quad (j, k = 1, 2, \ldots, n) \tag{A.8}$$

where $\det[\alpha_{jk}]$ is the determinant whose components are the quantities α_{jk}.

Note that within dominant publications promoted by the UR–QMS philosophy the interpretation of many-observable relations (A.8) is frequently omitted. The omission is due most probably to the fact that the idea of referring to simultaneous measurements for more than two observables is not supported convincingly by the current practice of experimental physics.

Addendum

Sometimes, in QM practice, a wave function Ψ is represented as a superposition of the form

$$\Psi = \sum_n \alpha_n \cdot \varphi_n, \qquad \sum_n |\alpha_n|^2 = 1, \tag{A.9}$$

were $\{\varphi_n\}$ denote a complete set of orthonormal basic functions for which $(\varphi_n, \varphi_m) = \delta_{nm}$ = a Kronecker delta.

Then, in a state described by Ψ, the mean value of an observable A is written as

$$\langle A \rangle_\Psi = \sum_{n,m} \alpha_n^* \cdot A_{nm} \cdot \alpha_m, \qquad A_{nm} = \left(\varphi_n, \hat{A}\,\varphi_m\right), \tag{A.10}$$

with A_{nm} indicating the matrix elements of operator $\hat{A}$ in representation given by $\{\varphi_n\}$.

When $\{\varphi_n\}$ are eigenfunctions of $\hat{A}$ the following formulas can be written

$$\hat{A}\,\varphi_n = a_n \cdot \varphi_n, \qquad \langle A\rangle_\Psi = \sum_n |\alpha_n|^2 \cdot a_n, \tag{A.11}$$

where a_n signify the eigenvalue of $\hat{A}$ in respect with the eigenfunction φ_n.

Note that the notations and formulas reminded in this Addendum can be used in connection with all quantities discussed above in present Appendix.

B: On the omission of conditions (A.3) within current literature

The mentioned omission is encountered in many generally agreed publications on QM (especially in textbooks, e.g. in [29]). It appears when the conventional Robertson-Schrödinger relation (A.7) is established by starting from the correct formula

$$\left\|\left(\left(\delta_\Psi\hat{A} + i\lambda\delta_\Psi\hat{B}\right)\Psi\right)\right\| \geqslant 0 \tag{B.1}$$

for the norm $\|f\|$ of function $f = (\delta_\Psi\hat{A} + i\lambda\delta_\Psi\hat{B})\Psi$. In (B.1) are used the notations presented in the previous Appendix A and λ denote a real and arbitrary parameter. In order to go on from this last formula to the relation (A.5), it is presumed the equality

$$\begin{aligned}&\left(\left(\delta_\Psi\hat{A} + i\lambda\delta_\Psi\hat{B}\right)\Psi, \left(\delta_\Psi\hat{A} + i\lambda\delta_\Psi\hat{B}\right)\Psi\right) = \\ &\left(\Psi, \left(\delta_\Psi\hat{A}\right)^2\Psi\right) + \lambda^2\left(\Psi, \left(\delta_\Psi\hat{B}\right)^2\Psi\right) \\ &-i\lambda\left(\Psi, \left[\hat{A}, \hat{B}\right]\Psi\right).\end{aligned} \tag{B.2}$$

Then, due to the fact that λ is a real and arbitrary quantity, from (B.1) it results the relation

$$\left\langle\left(\delta_\Psi\hat{A}\right)^2\right\rangle_\Psi \cdot \left\langle\left(\delta_\Psi\hat{B}\right)^2\right\rangle_\Psi \geqslant \frac{1}{4}\left|\left\langle\left[\hat{A}, \hat{B}\right]\right\rangle_\Psi\right|^2. \tag{B.3}$$

In terms of notations from Appendix A this last relation gives directly the formula

$$\Delta_\Psi A \cdot \Delta_\Psi B \geqslant \frac{1}{2}\left|\left\langle\left[\hat{A}, \hat{B}\right]\right\rangle_\Psi\right| \tag{B.4}$$

which is nothing but the relation (A.7) from the previous Appendix.

***Observation*:** Note here the next two aspects: (i) Introduction of (B.4) demands with necessity the existence of equality (B.2), (ii) The respective equality is true only when the operators $\hat{A}$ and $\hat{B}$ satisfy the conditions (A.3). The noted aspects must be signalized as omissions of the current literature.

Another context in which appears the omission of conditions (A.3) is connected with the "*braket notation*" frequently used in QM literature. Within the respective notation, known also as *Dirac's Notation* (DN), the scalar (inner) product of two functions f and g is depicted as $\langle f \mid g\rangle$ (see [29–31]). Of course DN was used in many texts regarding UR–QMS philosophy. But it must be pointed out the fact that in those texts the condition (A.3), justified in the previous Appendix, is totally omitted and its implications are not analyzed at all. It is easy to notice that such an omission is due to the fact that, within the DN, both terms (from left-hand and right-hand sides) of the condition (A.3) have the same transcription, namely:

$$\begin{aligned}&\left(\hat{A}_j\Psi, \hat{A}_k\Psi\right) = \left\langle\Psi\left|\hat{A}_j\hat{A}_k\right|\Psi\right\rangle \\ &\text{and} \\ &\left(\Psi, \hat{A}_j\hat{A}_k\Psi\right) = \left\langle\Psi\left|\hat{A}_j\hat{A}_k\right|\Psi\right\rangle.\end{aligned} \tag{B.5}$$

Obviously, such transcriptions create confusion and obstruct the just consideration of the condition (A.3) for cases where it is absolutely necessary in debates about UR–QMS philosophy. In order to avoid the above mentioned confusion in [32] we suggested that DN to be replaced by an Improved Dirac Notation (IDN). For such an IDN we proposed, that within scalar product of two functions f and g, to insert additionally the symbol "•" so that the respective product to be depicted as $< f \mid \bullet \mid g >$. In such a way it becomes directly visible the separation of the entities implied in that product. Then, inside of IDN, the two terms from (A.3) are transcribed as

$$\begin{aligned}&\left(\hat{A}_j\Psi, \hat{A}_k\Psi\right) = \left\langle\Psi\left|\hat{A}_j \bullet \hat{A}_k\right|\Psi\right\rangle \\ &\text{and} \\ &\left(\Psi, \hat{A}_j\hat{A}_k\Psi\right) = \left\langle\Psi\left|\bullet\hat{A}_j\hat{A}_k\right|\Psi\right\rangle\end{aligned} \tag{B.6}$$

Now one observes that in terms of IDN the condition (A.3) appears in the form

$$\text{iff} \qquad \langle\Psi|\,\hat{A}_j \bullet \hat{A}_k\,|\Psi\rangle = \langle\Psi|\bullet\,\hat{A}_j\hat{A}_k\,|\Psi\rangle \tag{B.7}$$

which no longer generates confusions in discussions about UR–QMS philosophy.

C: Classical "uncertainty relations" in Fourier analysis

In classical mathematical harmonic analysis it is known a relation (often named theorem) which, in terms of here used notations, is similar with the quantum UR depicted by relation (2). Through current mathematical representations the respective relation can be introduced as follows.

Let be a pair of variables x and ξ, with domains $x \in (-\infty, +\infty)$ and $\xi \in (-\infty, +\infty)$, regarded as arguments of a function $f(x)$ respectively of its Fourier transform

$$\tilde{f}(\xi) = \int_{-\infty}^{+\infty} \exp(-2i\pi\xi \cdot x) \cdot f(x) \cdot dx. \tag{C.1}$$

If the norm $\left\|f\right\|$ of $f(x)$ has the property $\left\|f\right\| = 1$, both $\left|f(x)\right|^2$ and $\left|\widetilde{f}(\xi)\right|^2$ are probability density functions for x and ξ regarded as real random (stochastic) variables. The variances of such variables, evaluated through the corresponding probabilities, can be noted as $\langle(x - \langle x\rangle^2)\rangle$ and $\langle(\xi - \langle\xi\rangle^2)\rangle$. The respective variances express the effective widths of functions $f(x)$ and $\widetilde{f}(\xi)$. Then [66] the aforesaid relation/theorem is given by the formula

$$\left\langle\left(x - \langle x\rangle^2\right)\right\rangle \cdot \left\langle\left(\xi - \langle\xi\rangle^2\right)\right\rangle \geqslant \frac{1}{16\pi^2}. \tag{C.2}$$

In mathematics this formula express the fact that: "*A nonzero function and its Fourier transform cannot both be sharply localized*" [66].

Often formula (C.2) is transcribed in a equivalent variant as follows

$$\Delta x \cdot \Delta\xi \geqslant \frac{1}{4\pi} \tag{C.3}$$

where Δx and $\Delta\xi$ denote the corresponding standard deviations of x and ξ, defined through conventions like $\Delta x = \sqrt{\langle(x - \langle x\rangle^2)\rangle}$. In non-quantum physics a version of relation (C.3) appears in studies of classical signals (waves of acoustic or electromagnetic nature) where $x = t = time$ and $\xi = \nu = frequency$. The respective version is written as

$$\Delta t \cdot \Delta\nu \geqslant \frac{1}{4\pi} \tag{C.4}$$

and it is known [67] as "*Gabor's uncertainty relation*". This last relation (C.4) means the fact that, for a classical signal (regarded as a wave packet), the product of the "uncertainties" ("irresolutions") Δt and $\Delta\nu$ in time and frequency domains cannot be smaller than a specific constant.

Formally the classical relation (C.3) can be transposed to the case of "quantum wave packets" often discussed in introductory/intuitive texts about QM. Such a transposition focuses on the pairs of conjugated observables *q–p* (coordinate-momentum) respectively *t–E* (time-energy). The corresponding transpositions can be obtained by setting in (C.4) the substitutions $x = q$ and $\xi = p\,(2\pi\hbar)^{-1}$ respectively $x = t$ and $\xi = E\,(2\pi\hbar)^{-1}$. The substitutions of variable ξ are nothing but the so called duality relations (regarding the wave-particle connections). By means of the mentioned substitutions from (C.4) one finds the following two relations

$$\Delta q \cdot \Delta p \geqslant \frac{\hbar}{2} \qquad \text{respectively} \qquad \Delta t \cdot \Delta E \geqslant \frac{\hbar}{2}. \tag{C.5}$$

These last formulas are similar with the conventional UR (2) for the pairs of observables *q–p* respectively *t–E*. Note that the mentioned similarity is admissible iff (if and only if) one accepts the conventions $\left|\langle[\hat{q}, \hat{p}]\rangle_\Psi\right| = \hbar$ and $\left|\langle[\hat{t}, \hat{E}]\rangle_\Psi\right| = \hbar$. But attention, the last convention has no more than a "metaphoric" value. This because in usual QM framework the time t is a deterministic but not random (stochastic) variable and, genuinely, for the respective framework a *time operator* $\hat{t}$ is nothing but a senseless and fictitious concept (see also the discussions from the deficiency $\boldsymbol{D_8}$).

Note that the classical relation (C.3) can be transposed also in another quantum formula regarding the ground state of a Quantum Torsion Pendulum (QTP) (see Subsection 3.6.2). For respective transposition in (C.3) it should to take $f(x) = \Psi(\varphi)$, $x = \varphi$ and $\xi = L_z \cdot (2\pi\hbar)^{-1}$. So one obtains the formula

$$\Delta\varphi \cdot \Delta L_z \geqslant \frac{\hbar}{2} \tag{C.6}$$

which is nothing but the lowest level version of the last of formulas (13)

Addendum

It is worth to mention here the fact that, in the Fourier analysis, the x-unlimited relations (C.3) and (C.4) have correspondent formulas in x-limited cases (when the variable x has a finite domain of existence). The respective fact can be evidenced as follows.

Let be $x \in [0, b)$, with b a finite quantity and function $f(x)$ having the property $f(0) = f(b - 0) := \lim\limits_{x \to b-0} f(x)$. Then the quantities

$$c_n = \frac{1}{\sqrt{b}} \int_0^b \exp(-ik_n x) \cdot f(x) \cdot dx \tag{C.7}$$

represent the Fourier coefficients of $f(x)$, with $k_n = n \cdot \frac{2\pi}{b}$ and n denoting integers i.e. $n \in \mathbb{Z}$.

Moreover if the measure $|f(x)|^2\,dx$ denotes the infinitesimal probability for $x \in (x, x + dx)$ the quantity $|c_n|^2$ signify the discrete probability associated to the value k_n. Then for functions $A = A(x)$ and $B = B(k_n)$, depending on x respectively on k_n, the mean (expected) values $\langle A\rangle$ and $\langle B\rangle$ are writen as follows

$$\begin{gathered} \langle A\rangle = \int_0^b A(x) \cdot |f(x)|^2\,dx, \\ \langle B\rangle = \sum_n B(k_n) \cdot |c_n|^2\,. \end{gathered} \tag{C.8}$$

As the most used such mean (expected) values can be quoted the following ones: first order moments $\langle x\rangle$ and $\langle k_n\rangle = \langle k\rangle$, variances $\langle(x - \langle x\rangle)^2\rangle$ and $\langle(k_n - \langle k\rangle)^2\rangle$ respectively standard deviations $\Delta x = \sqrt{\langle(x - \langle x\rangle)^2\rangle}$ and $\Delta k = \sqrt{\langle(k_n - \langle k\rangle)^2\rangle}$.

In order to find the announced x-limited correspondents of x-unlimited relations (C.3) and (C.4) we take into account the following obvious formula

$$\int_0^b \left|\lambda\,(x - \langle x\rangle) \cdot f(x) + \left(\frac{d}{dx} - i\,\langle k\rangle\right) \cdot f(x)\right|^2 \cdot dx \geqslant 0 \tag{C.9}$$

where λ is a real, finite and arbitrary parameter. By using the above noted probabilistic properties of function $f(x)$ and coefficients c_n from (C.9) one obtains the relation

$$\lambda^2 \left\langle(x - \langle x\rangle)^2\right\rangle + \lambda\left(b\,|f(0)|^2 - 1\right) + \left\langle(k - \langle k\rangle)^2\right\rangle \geqslant 0. \tag{C.10}$$

Due to the mentioned characteristics of λ, from this last relation one finds the next formulas for variances of x and k_n

$$\left\langle (x-\langle x\rangle)^2\right\rangle \cdot \left\langle (k_n-\langle k\rangle)^2\right\rangle \geqslant \frac{1}{4}\left(b\,|f(0)|^2-1\right)^2 \tag{C.11}$$

respectively for standard deviations of x and k_n

$$\Delta\mathrm{x}\cdot\Delta\mathrm{k} \geqslant \frac{1}{2}\left|\left(b\,|f(0)|^2-1\right)\right| . \tag{C.12}$$

The formulas (C.11) and (C.12) are x-limited analogues of the x-unlimited relations (C.2) and (C.3).

In the end we note that formula (C.12) is applicable in cases of wave functions (4) regarding non-degenerate circular rotations. For such cases the application of (C.12) is obtained through the following substitutions: $x \to \varphi$, $b \to 2\pi$, $f(x) \to \Psi(\varphi)$ and $k_n \to \frac{L_z}{\hbar}$. So from (C.12) it results

$$\Delta\varphi\cdot\Delta L_z \geqslant \frac{\hbar}{2}\left|\left(2\pi\,|\Psi(0)|^2-1\right)\right| . \tag{C.13}$$

This last formula in case of wave functions (4) degenerates into trivial equality $0 = 0$

D: On the fluctuations of thermodynamic observables

Thermodynamic systems are macroscopic bodies composed by huge numbers of microscopic constituents (molecules and atoms). As whole bodies or through by their macroscopic parts such systems are described by so-called thermodynamic observables. The alluded observables are viewed as deterministic variables (in usual thermodynamics) respectively as stochastic quantities (in statistical physics). In the last view they are characterized by fluctuations (deviations from their deterministic values studied within usual thermodynamics). The mentioned fluctuations are investigated within the next conceptual frameworks: (a) phenomenological approach, (b) classical statistical mechanics, respectively (c) quantum statistical physics.

In phenomenological approach [68–72], proposed for the first time by Einstein, the respective fluctuations can be depicted briefly as follows. Let be a system of the mentioned kind, whose properties are described by a set of thermodynamic observables $\mathbb{A}_j$ $(j = 1, 2, 3, \ldots, n)$. Each such observable $\mathbb{A}_j$ is characterized by a global fixed value $\overline{\mathbb{A}}_j$, evaluable through the methods of deterministic usual thermodynamics. Then the fluctuations of observables $\mathbb{A}_j$ should be discussed in terms of random variables $\mathcal{A}_j = \mathbb{A}_j - \overline{\mathbb{A}}_j$ $(j = 1, 2, \ldots, n)$, endowed with continuous spectra of values such are $\mathcal{A}_j \in (-\infty, +\infty)$. The random characteristics of variables $\mathcal{A}_j$, i.e. the fluctuations of observables $\mathbb{A}_j$, are depicted in phenomenological approach through the probability density $W = W(\vec{\mathcal{A}})$, where the vector $\vec{\mathcal{A}}$ signifies the set of all variables $\mathcal{A}_j$. Commonly for $W = W(\vec{\mathcal{A}})$ one uses distributions of Gaussian type. The mean value (expected) value $\langle\mathbb{A}_j\rangle_W$ and the random deviation $\delta_W\mathbb{A}_j$ of the observable $\mathbb{A}_j$ are

$$\begin{gathered}\left\langle\mathbb{A}_j\right\rangle_W = \int_{-\infty}^{+\infty} \mathbb{A}_j\cdot W\left(\vec{\mathcal{A}}\right)\cdot d\vec{\mathcal{A}},\\ \delta_W\mathbb{A}_j = \mathbb{A}_j - \left\langle\overline{\mathbb{A}}_j\right\rangle_W = \mathcal{A}_j.\end{gathered} \tag{D.1}$$

Usually, the fluctuations of observables $\mathbb{A}_j$ $(j = 1, 2, 3, \ldots, n)$ are characterized by a small number of numerical parameters evaluable through the random deviations $\delta_W\mathbb{A}_j$. Examples of such parameters are: dispersions $\langle(\delta_W\mathbb{A}_j)^2\rangle_W = \langle(\mathcal{A}_j)^2\rangle_W$ and their equivalents the standard deviations $\Delta_W\mathbb{A}_j = \sqrt{\langle(\delta_W\mathbb{A}_j)^2\rangle_W}$, second order moments (correlations) $\langle\delta_W\mathbb{A}_j\cdot\delta_W\mathbb{A}_k\rangle_W$ $(j \neq k)$ or even [72] higher order moments (correlations) $\langle(\delta_W\mathbb{A}_j)^r\cdot(\delta_W\mathbb{A}_k)^s\rangle_W$, $(r + s \geqslant 3)$.

The correlations $\langle\delta_W\mathbb{A}_j\cdot\delta_W\mathbb{A}_k\rangle_W$ $(j, k = 1, 2, \ldots, n)$ constitute the components of a positive semi definite matrix. The respective components satisfy [70, 71] the following correlation formulas

$$\det\left[\left\langle\delta_W\mathbb{A}_j\cdot\delta_W\mathbb{A}_k\right\rangle_W\right] \geqslant 0\,, \tag{D.2}$$

where $\det[\alpha_{jk}]$ denote the determinant whose components are the quantities α_{jk}. Particularly for two thermodynamic observables $\mathbb{A}_1 = \mathbb{A}$ and $\mathbb{A}_2 = \mathbb{B}$ from (D.2) one obtains

$$\Delta_W\mathbb{A}\cdot\Delta_W\mathbb{B} \geqslant |\langle\delta_W\mathbb{A}\cdot\delta_W\mathbb{B}\rangle_W| \tag{D.3}$$

where $\Delta_W A = \sqrt{\langle(\delta_W A)^2\rangle_W}$ denotes the standard deviation of observable $\mathbb{A}$. Mathematically (in sense of probability theory) this last classical formula is completely analogous with the quantum UR (2).

Regarded in their detailed expressions the standard deviations like is $\Delta_W\mathbb{A}$ (introduced above) have an interesting generic property. Namely they appear as being in a direct and factorized dependence of Boltzmann's constant k_B. The respective dependence has the following physical significance. It is known the fact that, mathematically, for a given quantity the standard deviation indicates its randomness. This in the sense that the respective quantity is a random or, alternatively, a deterministic (non-random) variable according as the alluded deviation has a positive or null value. Therefore $\Delta_W\mathbb{A}$ can be regarded as an indicator of randomness for the thermodynamic observable $\mathbb{A}$. But, for diverse cases (of observables, systems and states), the deviation $\Delta_W\mathbb{A}$ has various expressions in which, apparently, no common element seems to be implied. Nevertheless such an element can be found out [20,73] as being materialized by the Boltzmann's constant k_B. So, in Gaussian approximation within the framework of phenomenological theory of fluctuations one finds [20, 73]

$$(\Delta_W\mathbb{A})^2 = k_B\cdot\sum_\alpha\sum_\beta\frac{\partial\bar{\mathbb{A}}}{\partial\bar{\mathbb{X}}_\alpha}\cdot\frac{\partial\bar{\mathbb{A}}}{\partial\bar{\mathbb{X}}_\beta}\cdot\left(\frac{\partial^2\bar{\mathbb{S}}}{\partial\bar{\mathbb{X}}_\alpha\partial\bar{\mathbb{X}}_\beta}\right)^{-1} . \tag{D.4}$$

Now note that, a kind of non-quantum formulas completely similar with (D.2) and (D.3), can be reported also for the fluctuations of thermodynamic observables described in terms of classical statistical mechanics. In the respective terms the above phenomenological notations and relations can be transcribed formally as follows. Instead of random variables $\mathcal{A}_j$ should to operate with the phase space ensemble denoted as μ of all coordinates and momenta of molecules/atoms which compose the thermodynamic system. Also instead of observables $\mathbb{A}_j = \overline{\mathbb{A}}_j + \mathcal{A}_j$ needs to be use the random functions of the form $\mathbb{A}_j = \mathbb{A}_j(\mu)$. Therewith the probability density $W = W(\vec{\mathcal{A}})$ should to be replaced with the statistical distribution function $w = w(\mu)$. Then, in terms of aforesaid description of considered fluctuations, as example, can be written the relation

$$\Delta_w \mathbb{A} \cdot \Delta_w \mathbb{B} \geqslant \left| \langle \delta_w \mathbb{A} \cdot \delta_w \mathbb{B} \rangle_w \right| \tag{D.5}$$

which is completely similar with (D.3). Add here the observation that the standard deviations $\Delta_w \mathbb{A}$ and $\Delta_w \mathbb{B}$ from (D.5) have a factorization dependence on k_B of type (D.4), similarly with the case of quantities $\Delta_W \mathbb{A}$ and $\Delta_W \mathbb{B}$ from (D.3).

For describing the fluctuations of thermodynamic observables $\mathbb{A}_j$ in framework of quantum statistical physics as probabilities carrier instead of phenomenological density $W = W(\vec{A})$ should to use [73–76] the quantum density operator $\hat{\rho}$:

$$\hat{\rho} = \sum_k \mathfrak{p}_k \left| \psi_k \right\rangle \left\langle \psi_k \right| . \tag{D.6}$$

Here $\left| \psi_k \right\rangle$ $(k = 1, 2, \dots)$ denote the wave functions of pure states of system and $\mathfrak{p}_k$ are the corresponding probabilities of the respective states. In the same framework the above mentioned random variables $\mathcal{A}_j$ are substituted with the thermo-quantum operators $\hat{\mathbb{A}}_j$ $(j = 1, 2, \dots, n)$. In framework of quantum statistical physics the mean value $\langle \mathbb{A}_j \rangle_\rho$ and random deviation $\delta_\rho \hat{\mathbb{A}}_j$ of observable $\mathbb{A}_j$ are

$$\begin{aligned} \left\langle \mathbb{A}_j \right\rangle_\rho &= \sum_k \mathfrak{p}_k \left\langle \psi_k \right| \hat{\mathbb{A}}_j \left| \psi_k \right\rangle \\ &= tr\left(\sum_k \mathfrak{p}_k \left| \psi_k \right\rangle \left\langle \psi_k \right| \hat{\mathbb{A}}_j \right) = tr\left(\hat{\rho} \cdot \hat{\mathbb{A}}_j \right), \\ \delta_\rho \hat{\mathbb{A}}_j &= \hat{\mathbb{A}}_j - \left\langle \mathbb{A}_j \right\rangle_\rho . \end{aligned} \tag{D.7}$$

The deviations $\delta_\rho \hat{\mathbb{A}}_j$ can be used in description of numerical parameters of fluctuations for observables $\mathbb{A}_j$ in the mentioned framework. As such parameters can be quoted: dispersions $\langle (\delta_\rho \hat{\mathbb{A}}_j)^2 \rangle_\rho$ and their equivalents standard deviations $\Delta_\rho \mathbb{A}_j = \sqrt{\langle (\delta_\rho \hat{\mathbb{A}}_j)^2 \rangle_\rho}$, second order moments (correlations) $\langle \delta_\rho \hat{A}_j \cdot \delta_\rho \hat{A}_k \rangle_\rho$ (where $j \neq k$) or even higher order moments $\langle (\delta_\rho \hat{A}_j)^r \cdot (\delta_\rho \hat{A}_k)^s \rangle_\rho$ (where $r + s \geqslant 3$).

In case of two thermodynamic observables $\mathbb{A}$ and $\mathbb{B}$, regarded in framework of quantum statistical physics, can be introduced also a correlation relation similar with (D.3) and (D.5). Such a relation can be introduced as follows. For the corresponding thermo-quantum operators $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$ it is evidently true the relation

$$\sum_k \mathfrak{p}_k \left\langle \left(\delta_\rho \hat{\mathbb{A}} + i\lambda\, \delta_\rho \hat{\mathbb{B}} \right) \psi_k \;\middle|\; \left(\delta_\rho \hat{\mathbb{A}} + i\lambda\, \delta_\rho \hat{\mathbb{B}} \right) \psi_k \right\rangle \geqslant 0 \tag{D.8}$$

where λ is an arbitrary real parameter. If in respect with the functions ψ_k the operators $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$ satisfy the conditions of type (A.3) one can write

$$\begin{aligned} &\sum_k \mathfrak{p}_k \left\langle \left(\delta_\rho \hat{\mathbb{A}} + i\lambda\, \delta_\rho \hat{\mathbb{B}} \right) \psi_k \;\middle|\; \left(\delta_\rho \hat{\mathbb{A}} + i\lambda\, \delta_\rho \hat{\mathbb{B}} \right) \psi_k \right\rangle \\ &= \sum_k \mathfrak{p}_k \left\langle \psi_k \middle| \left(\delta_\rho \hat{\mathbb{A}} \right)^2 \middle| \psi_k \right\rangle \\ &\quad + \lambda^2 \sum_k \mathfrak{p}_k \left\langle \psi_k \middle| \left(\delta_\rho \hat{\mathbb{B}} \right)^2 \middle| \psi_k \right\rangle \\ &\quad + i\lambda \sum_k \mathfrak{p}_k \left\langle \psi_k \middle| \left(\delta_\rho \hat{\mathbb{A}} \cdot \delta_\rho \hat{\mathbb{B}} - \delta_\rho \hat{\mathbb{B}} \cdot \delta_\rho \hat{\mathbb{A}} \right) \middle| \psi_k \right\rangle . \end{aligned} \tag{D.9}$$

Then from (D.8) it results the relation

$$\left\langle \left(\delta_\rho \hat{\mathbb{A}} \right)^2 \right\rangle_\rho + \lambda^2 \left\langle \left(\delta_\rho \hat{\mathbb{B}} \right)^2 \right\rangle_\rho + \lambda \left\langle i \left[\hat{\mathbb{A}}, \hat{\mathbb{B}} \right] \right\rangle_\rho \geqslant 0 \tag{D.10}$$

where $[\hat{\mathbb{A}}, \hat{\mathbb{B}}]$ denotes the commutator of operators $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$.

Because λ is an arbitrary real parameter from (D.10) one obtains the relation

$$\left\langle \left(\delta_\rho \hat{\mathbb{A}} \right)^2 \right\rangle_\rho \cdot \left\langle \left(\delta_\rho \hat{\mathbb{B}} \right)^2 \right\rangle_\rho \geqslant \frac{1}{4} \left\langle i \left[\hat{\mathbb{A}}, \hat{\mathbb{B}} \right] \right\rangle_\rho^2 \tag{D.11}$$

or the equivalent formula

$$\Delta_\rho \mathbb{A} \cdot \Delta_\rho \mathbb{B} \geqslant \frac{1}{2} \left| \left\langle \left[\hat{\mathbb{A}}, \hat{\mathbb{B}} \right] \right\rangle_\rho \right| . \tag{D.12}$$

Now let us remind the fact that in quantum statistics the above discussed thermo-quantum quantities $\langle (\delta_\rho \hat{\mathbb{A}}_j)^2 \rangle_\rho$ and $\Delta_\rho A$ are proved to be connected directly with a quantity from deterministic (simple thermodynamic) description of thermodynamic observables. The respective connection is due by the known fluctuation-dissipation theorem [76] which is expressed by the relation

$$\begin{aligned} \left\langle \left(\delta_\rho \hat{\mathbb{A}} \right)^2 \right\rangle_\rho &= \left(\Delta_\rho \mathbb{A} \right)^2 \\ &= \frac{\hbar}{2\pi} \int_{-\infty}^{+\infty} \coth\left(\frac{\hbar\omega}{2k_B T} \right) \cdot \mathcal{X}''(\omega) \cdot d\omega . \end{aligned} \tag{D.13}$$

Here k_B = the Boltzmann's constant, $\hbar$ = Planck's constant and T = temperature of the considered system. Also in (D.13) the quantity $\mathcal{X}''(\omega)$ denote the imaginary part of the susceptibility associated with the observable $\mathbb{A}$. Note that $\mathcal{X}''(\omega)$ is a deterministic quantity which is defined primarily in non-stochastic framework of macroscopic physics [77]. Due to the respective definition it is completely independent of both k_B and $\hbar$.

In the end of this Appendix the following conclusion may be recorded. All the relations (D.2), (D.3), (D.4), (D.10) and (D.11) are formulas regarding macroscopic fluctuations but not pieces which should be adapted to the UR–QMS philosophy requirements.

E: On the measurements of macroscopic fluctuations

The fluctuations parameters, defined above Appendix D, refer to the characteristics of intrinsic nature for the considered macroscopic systems. But in practical actions, for the same systems, one operates with global parameters, of double source (origin). A first source is given by the intrinsic properties of systems. A second source is provided by the actions of measuring devices. In such a vision a measurement can be regarded as an transmission process of information (refering to stochastic data). Consequently the data about the intrinsic properties of measured system appear as *input* (*in*) information while the global results of the corresponding measurement represent the *output* (*out*) information.

Here below we will appeal to the aforesaid vision for giving (as in [91, 107]) a theoretical model regarding the measurement of thermal fluctuations. The respective fluctuations will be considered in a phenomenological approach (see Appendix D). For simplicity let us consider a system characterized by a single macroscopic observable $\mathbb{A} = \overline{\mathbb{A}} - \mathcal{A}$, whose thermal fluctuations are impacted within the random variable $\mathcal{A}$ having the spectrum $\mathcal{A} \in (-\infty, +\infty)$. The intrinsic fluctuations of $\mathcal{A}$ is supposed to be described by the probability distribution $W_{in} = W_{in}(\mathcal{A})$ regarded as carrier of input-information. The results of measurements are depicted by the distribution $W_{out} = W_{out}(\mathcal{A})$ regarded as bearer of out-information. Then the measuring process may be symbolized as a transformation of the form $W_{in}(\mathcal{A}) \rightarrow W_{out}(\mathcal{A})$. If the measuring device is supposed to have stationary and linear characteristics, the mentioned transformation can be described as follows:

$$W_{out}(\mathcal{A}) = \int_{-\infty}^{+\infty} K(\mathcal{A}, \mathcal{A}') \cdot W_{in}(\mathcal{A}') \cdot d\mathcal{A}' \tag{E.1}$$

where $K(\mathcal{A}, \mathcal{A}')$ appears as a doubly stochastic kernel (in sense defined in [108]). This means that $K(\mathcal{A}, \mathcal{A}')$ satisfy the relations $\int_{-\infty}^{+\infty} K(\mathcal{A}, \mathcal{A}')\, d\mathcal{A} = \int_{-\infty}^{+\infty} K(\mathcal{A}, \mathcal{A}')\, d\mathcal{A}' = 1$.

Add here the fact that, from a physical perspective, the kernel $K(\mathcal{A}, \mathcal{A}')$ incorporates the theoretical description of all the characteristics of the measuring device. Particularly, for an ideal device which ensure $W_{out}(\mathcal{A}) = W_{in}(\mathcal{A})$, it must to have the expression $K(\mathcal{A}, \mathcal{A}') = \delta(\mathcal{A} - \mathcal{A}')$, where $\delta(\mathcal{X})$ denote the Dirac's δ-function of argument $\mathcal{X}$.

By means of distributions $W_\eta = W_\eta(\mathcal{A})$ ($\eta = in; out$) can be introduced the corresponding η-numerical-characteristics of thermal fluctuations of observable $\mathbb{A} = \overline{\mathbb{A}} + \mathcal{A}$. Such are the η-mean (expected) value $\langle \mathbb{A} \rangle_\eta$ and η - standard deviation $\Delta_\eta \mathbb{A}$ defined through the relations

$$\begin{aligned} \langle \mathcal{A} \rangle_\eta &= \int_{-\infty}^{+\infty} \mathcal{A} \cdot W_\eta(\mathcal{A}) \cdot d\mathcal{A}, \\ \left(\Delta_\eta \mathcal{A}\right)^2 &= \left\langle \left(\mathcal{A} - \langle \mathcal{A} \rangle_\eta\right)^2 \right\rangle_\eta . \end{aligned} \tag{E.2}$$

The above considerations allow to note some observations about the measuring uncertainties (errors) regarding the fluctuating macroscopic observable $\mathbb{A}$. Firstly the $\eta = in$-versions of the parameters (E.2) describe only the "intrinsic" properties of the measured system. Secondly the $\eta = out$-variants of the same parameters incorporate composite information about the respective system and considered measuring device. That is why one can say that, in terms of the above discussions, the measuring uncertainties of observable $\mathbb{A}$ should be described by the following error indicators (characteristics)

$$\begin{aligned} \mathcal{E}\{\langle \mathbb{A} \rangle\} &= \langle \mathbb{A} \rangle_{out} - \langle \mathbb{A} \rangle_{in} , \\ \mathcal{E}\{\Delta\, \mathbb{A}\} &= \Delta_{out}\, \mathbb{A} - \Delta_{in}\, \mathbb{A}. \end{aligned} \tag{E.3}$$

Observe here that because $\mathbb{A}$ has stochastic characteristics for a relevant description of its measuring uncertainties it is completely insufficient the single indicator $\mathcal{E}\{\langle \mathbb{A} \rangle\}$. An adequate minimal such description requires at least the couple $\mathcal{E}\{\langle \mathbb{A} \rangle\}$ and $\mathcal{E}\{\Delta\, \mathbb{A}\}$. For further approximations of errors caused by measurements can be taken into account [111] the higher order moments like the next ones

$$\mathcal{E}\{\langle (\delta \mathbb{A})^n \rangle\} = \langle (\delta_{out} \mathbb{A})^n \rangle_{out} - \langle (\delta_{in} \mathbb{A})^n \rangle_{in} \tag{E.4}$$

where $\delta_\eta \mathbb{A} = \mathbb{A} - \langle \mathbb{A} \rangle$, $\eta = in, out$ and $n \geqslant 3$.

F: An exemplification for subsection 5.2

For presenting the announced exemplification we will refer to QMS of the energy for a particle of mass m, located in an infinite square well potential of width L [29]. The intrinsic state of the microparticle will be considered as being described by the *in*-wave function $\Psi_{in}(x) = \sum_{j=1}^{n} \alpha_j \cdot \varphi_j(x)$. Here $\varphi_j(x)$ denote the eigenfunctions associated to the energetic eigenvalues $a_j = E_j = \Im \cdot j^2$ where $\Im = \left(\hbar^2 \pi^2 / 2mL^2\right)$ and $j = 1, 2, 3, \ldots$. In the considered *in*-wave function the quantities α_j are probability amplitudes corresponding to the eigenvalues E_j.

We will restrict our exemplification by taking into account only the following circumstances. So we take $n = 3$ as the upper value of the inner energy of the particle while for the amplitudes α_j we will consider the values which give $(|\alpha_j|^2) = (0.5 \;\; 0.4 \;\; 0.1)$.

Then the intrinsic characteristics of the particle energy are described by the next mean value and the standard deviation

$$\langle E \rangle_{in} = 3 \cdot \Im, \qquad \Delta_{in} E = 2.45 \cdot \Im. \tag{F.1}$$

Accordingly with discussions from Subsection 5.2, for a particle in the mentioned intrinsic state, the measurement of energy can be described as follows. We need to define a model-expression for the matrix (M_{kj}) from (29). As a first example, we will consider a measurement done with a device endowed with flawed (FL) characteristics. Such devices, for instance, can be associated with a matrix (M_{kj}) having the form

$$\left(M_{kj}\right)_{FL} = \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.4 & 0.2 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}. \tag{F.2}$$

Thus the outcomes of measurement will be characterized by probabilities $\left(|\beta_k|^2\right)_{FL} = (0.34 \ \ 0.38 \ \ 0.23)$. With such probabilities, the measurement outcomes for energy will be characterized by the next FL-expected-value and FL-standard-deviation

$$\langle E\rangle_{FL} = 3.98 \cdot \Im, \qquad \Delta_{FL} E = 3.04 \cdot \Im. \tag{F.3}$$

Consequently, for the measurement described by (F.2), the error indicators (29) and (31) acquire the following FL-values

$$\mathcal{E}_{FL}\{\langle E\rangle\} = 0.9 \cdot \Im, \qquad \mathcal{E}_{FL}\{\Delta E\} = 0.59 \cdot \Im. \tag{F.4}$$

If, for the above mentioned energy/particle, we want to describe a measurement done with a device having larger characteristics of accuracy (ACC) one can proceed as follows. In the spirit of the relations (31), for the matrix (M_{kj}) instead of the formula (F.2) we appeal, for example, to the following expression

$$\left(M_{kj}\right)_{ACC} = \begin{pmatrix} 0.95 & 0.03 & 0.02 \\ -0.03 & 1.04 & -0.01 \\ 0.08 & -0.07 & 0.99 \end{pmatrix}. \tag{F.5}$$

So, for the probabilities associated to the outcomes of ACC-measurement, one obtains $(|\beta_k|^2)_{ACC} = (0.489 \ \ 0.4 \ \ 0.11)$. Associated to the respective probabilities, the considered ACC-measurement of energy is characterized by the next ACC-expected value and ACC-standard-deviation

$$\langle E\rangle_{ACC} = 3.088 \cdot \Im, \qquad \Delta_{ACC} E = 2.52 \cdot \Im. \tag{F.6}$$

By comparing values from (F.6) with those from (F.1) one sees that the referred ACC-measurement is characterized by the following error indicators

$$\mathcal{E}_{ACC}\{\langle E\rangle\} = 0.08 \cdot \Im, \qquad \mathcal{E}_{ACC}\{\Delta E\} = 0.07 \cdot \Im. \tag{F.7}$$

Finally, by comparing the results reported in relations (F.4) and (F.7), we can note the following remark. Within the above theoretical description of measurement, the error indicators (for both expected value and standard deviation) are much smaller in the case dealing with higher accuracy characteristics comparatively with the one regarding flawed features.

G: Illustrations for subsection 5.4

In order to illustrate the model discussed in Subsection 5.4, in connection with the description of QMS, let us present here an exercise taken by abbreviation from our article [20] (more computational details can be found in the respective article). We will refer to a micro-particle of mass m having an one-dimensional motion along the x-axis. Its *in*-wave-function Ψ_{in} is taken of the form $\Psi_{in}(x) = |\Psi_{in}(x)| \cdot \exp\{i\Phi_{in}(x)\}$ where

$$|\Psi_{in}(x)| \propto \exp\left\{-\frac{(x-x_0)^2}{4\sigma^2}\right\}, \qquad \Phi_{in}(x) = kx. \tag{G.1}$$

Here as well as below in other relations from this Appendix the explicit notations of normalization constants are omitted (they can be added easy by the interested readers). According to the wave function (G.1) the intrinsic features of the considered microparticle are described by the parameters x_0, σ and k.

Through expressions (G.1), by means of formulas (38), it is simple to find the analytical expresions for probability density ρ_{in} and current j_{in}. As doubly stochastic kernels suggested in (40) we propose here the next two formulas

$$\Gamma(x, x') \propto \exp\left\{-\frac{(x-x')^2}{2\gamma^2}\right\}, \tag{G.2}$$

$$\Lambda(x, x') \propto \exp\left\{-\frac{(x-x')^2}{2\lambda^2}\right\}. \tag{G.3}$$

Here parameters γ and λ depict the characteristics of measuring devices/procedures. The values of the respective parameters are associated with an ideal measurement (when both γ and λ tend toward zero), respectively with a nonideal measurement (in cases when at least one of the two parameters is not-null).

Then, by using the procedures presented within Subsection 5.4, it is easy to find the *out*-entities ρ_{out}, j_{out} and Ψ_{out}. By using the respective entities together with the functions from (G.1) one can evaluate the *out* and *in* versions of mean (expected) values and standard deviations for observables of interest. The respective evaluations ensure estimations of the corresponding error indicators. So, for $\hat{x} = x\cdot$ = coordinate and $\hat{p} = -i\hbar\nabla_x$ = momentum as operators (observables) of interest, one obtains [20] the following error indicators

$$\mathcal{E}\{\langle x\rangle\} = 0, \quad \mathcal{E}\{\Delta x\} = \sqrt{\sigma^2+\gamma^2} - \sigma, \tag{G.4}$$

$$\begin{aligned} &\mathcal{E}\{\langle p\rangle\} = 0, \\ &\mathcal{E}\{\Delta p\} = \hbar \left| \left[\frac{k^2(\sigma^2+\gamma^2)}{\sqrt{(\sigma^2+\lambda^2)(\sigma^2+2\gamma^2-\lambda^2)}} - \right.\right. \\ &\qquad\qquad \left.\left. -k^2 + \frac{1}{4(\sigma^2+\gamma^2)} \right]^{\frac{1}{2}} - k \right|. \end{aligned} \tag{G.5}$$

Let us now restrict in the wave function (G.1) to the situation when $x_0 = 0$ $k = 0$ and $\sigma = \sqrt{\frac{\hbar}{2m\omega}}$. Then (G.1) describe the ground state of a harmonic oscillator with m = mass and ω = angular frequency. As observable of interest of such an oscillator we consider the energy described by the Hamiltonian

$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\,\omega^2}{2}x^2$. For the respective observable one finds

$$\langle H \rangle_{in} = \frac{\hbar\omega}{2}, \qquad \Delta_{in} H = 0, \tag{G.6}$$

$$\langle H \rangle_{out} = \frac{\omega\left[\hbar^2 + \left(\hbar + 2m\,\omega\,\gamma^2\right)^2\right]}{4\left(\hbar + 2m\,\omega\,\gamma^2\right)}, \tag{G.7}$$

$$\Delta_{out} H = \frac{\sqrt{2}\,m\,\omega^2\,\gamma^2\left(\hbar + m\,\omega\,\gamma^2\right)}{\left(\hbar + 2m\,\omega\,\gamma^2\right)}, \tag{G.8}$$

$$\mathcal{E}\{\langle H \rangle\} = \frac{m^2\omega^3\gamma^4}{\hbar + 2m\,\omega\,\gamma^2}, \tag{G.9}$$

$$\mathcal{E}\{\Delta H\} = \Delta_{out} H = \frac{\sqrt{2}\,m\,\omega^2\,\gamma^2\left(\hbar + m\,\omega\,\gamma^2\right)}{\left(\hbar + 2\,m\,\omega\,\gamma^2\right)}. \tag{G.10}$$

H: A more comprehensive description of measuring errors for random observables

In Subsections 5.2 and 5.4 or Appendices E, F and G, we have discussed the measuring errors for random observables of quantum respectively macroscopic nature. For description of that errors, were used as indicators only the lower order probabilistic parameters (moments and correlations). But those indicators give only first sequences, of limited value, for a global picture of the considered errors. A more comprehensive such a picture can be done in terms of informational entropies. Shortly, for the above discussed observables and errors, the suggested depiction can be illustrated as follows.

Firstly let us refer to the case of a macroscopic random observable $\mathcal{A}$ whose measurements are outlined in Appendix E. The intrinsic characteristics (fluctuations) of $\mathcal{A}$ are considered as being described by the probability distribution $W_{in} = W_{in}(\mathcal{A})$ regarded as carrier of input-information for measurements. The results of measurements are depicted by the distribution $W_{out} = W_{out}(\mathcal{A})$ associated with the out-information of measurements. The informational entropies $\mathcal{H}_\eta$ ($\eta = in, out$) connected with the above noted distributions are defined through the formulas

$$\mathcal{H}_\eta(\mathcal{A}) = -\int_{-\infty}^{+\infty} W_\eta(\mathcal{A}) \cdot \ln\left[W(\mathcal{A})\right] \cdot d\mathcal{A}. \tag{H.1}$$

By taking into account the transformation (E.1), the main properties of the doubly stochastic kernel $K(\mathcal{A}, \mathcal{A}')$, as well the formula $\ln(X) \leqslant X - 1$ one can write

$$\begin{aligned}
&\mathcal{H}_{out}(\mathcal{A}) - \mathcal{H}_{in}(\mathcal{A}) \\
&= -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} d\mathcal{A} \cdot d\mathcal{A}' \cdot K(\mathcal{A}, \mathcal{A}') \cdot W_{in}(\mathcal{A}') \\
&\qquad \cdot \ln\left[\frac{W_{out}(\mathcal{A})}{W_{in}(\mathcal{A}')}\right] \\
&\geqslant -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} d\mathcal{A} \cdot d\mathcal{A}' \cdot K(\mathcal{A}, \mathcal{A}') \cdot W_{in}(\mathcal{A}') \\
&\qquad \cdot \left[\frac{W_{out}(\mathcal{A})}{W_{in}(\mathcal{A}')} - 1\right] \\
&= 0.
\end{aligned} \tag{H.2}$$

Therefore the errors specific of measurements for $\mathcal{A}$ in its wholeness can be described through the comprehensive error indicator

$$\mathcal{E}\{\mathcal{H}(\mathcal{A})\} = \mathcal{H}_{out}(\mathcal{A}) - \mathcal{H}_{in}(\mathcal{A}) \geqslant 0. \tag{H.3}$$

This relationship shows that the measuring process can be described by a non-negative change in the informational entropy associated with the investigated observable. The situation when the respective change is null corresponds to the case of an ideal measurement (free of errors), mentioned otherwise in connection with the relationship (E.1).

Mostly, the macroscopic fluctuations described by the here used observable $\mathcal{A}$ are investigated in the so-called Gaussian approximations. Then the entities $W_{in}(\mathcal{A})$ and $K(\mathcal{A}, \mathcal{A}')$ which appear in (E.1) are given by the following formulas

$$\begin{aligned}
W_{in}(\mathcal{A}) &\propto \exp\left\{-\frac{\mathcal{A}^2}{2a^2}\right\}, \\
K(\mathcal{A},\ \mathcal{A}') &\propto \exp\left\{-\frac{(\mathcal{A} - \mathcal{A}')^2}{2b^2}\right\},
\end{aligned} \tag{H.4}$$

where the explicit indication of normalization constants are omitted (the omission can be filled easily by interested readers). In the first formula from (H.4) a denotes the standard deviation of intrinsic fluctuations within the measured system. The symbol b in the second expression from (H.4) depicts the precision parameter of measuring device. Of course, for a scientifically acceptable measuring process, it must be considered that $b \ll a$.

In the alluded cases with Gaussian approximations the output distribution $W_{out}(\mathcal{A})$ has the form

$$W_{out}(\mathcal{A}) \propto \exp\left\{-\frac{\mathcal{A}^2}{2\left(a^2 + b^2\right)}\right\}.$$

Then the comprehensive error indicator (H.3) becomes

$$\mathcal{E}\{\mathcal{H}(\mathcal{A})\} = \frac{1}{2}\ln\left(1 + \frac{b^2}{a^2}\right) \approx \frac{1}{2} \cdot \frac{b^2}{a^2}. \tag{H.5}$$

Now let us refer to the comprehensive informational depiction for measuring errors in cases of random quantum observables. We start the announced reference by discussing the case presented in Subsection 5.2, regarding the measurement of a quantum observable endowed with a discrete spectrum of eigenvalues. In the respective case the input and output data characterizing the measurement are depicted by the following corresponding probabilities

$$\mathcal{P}_{in}^{j} = \left|\alpha_j\right|^2 \ , \quad \mathcal{P}_{out}^{j} = \left|\beta_j\right|^2 \ , \qquad (j = 1, 2, \ldots, n)\,. \tag{H.6}$$

These probabilities can be associated with the next information entropies

$$\mathcal{H}\left(\mathcal{P}_\eta\right) = -\sum_{j=1}^{n} P_\eta^j \cdot \ln\left(\mathcal{P}_\eta^j\right) \ , \quad (\eta = in,\ out)\,. \tag{H.7}$$

Consequently, for an extensive description of measuring errors for the specified quantum observable, can be used the below comprehensive indicator

$$\mathcal{E}\{\mathcal{H}(\mathcal{P})\} = \mathcal{H}(\mathcal{P}_{out}) - \mathcal{H}(\mathcal{P}_{in})\,. \tag{H.8}$$

By taking into account the transformation (27), the basic properties of doubly stochastic matrix M_{jk}, plus the evident formula $\ln(X) \leqslant X - 1$, through some simple calculations (similar to those appealed in (H.2) and (H.3), one finds:

$$\mathcal{E}\{\mathcal{H}(\mathcal{P})\} \geqslant 0. \tag{H.9}$$

This formula corresponds to ideal or non-ideal measurements, in cases of equality respectively of inequality.

Note that, in cases of examples presented in Appendix F related with Subsection 5.2, the relation (H.8) takes the expresions

$$\begin{aligned} \mathcal{E}\{\mathcal{H}(\mathcal{P})_{FL}\} &= \mathcal{H}\Big(\left(|\beta_k|^2\right)_{FL}\Big) - \mathcal{H}\Big(\left(|\alpha_j|^2\right)\Big) \\ &= 0.131, \\ \mathcal{E}\{\mathcal{H}(\mathcal{P})_{ACC}\} &= \mathcal{H}\Big(\left(|\beta_k|^2\right)_{ACC}\Big) - \mathcal{H}\Big(\left(|\alpha_j|^2\right)\Big) \\ &= 0.018. \end{aligned} \tag{H.10}$$

The above expressions correspond to measurements with characteristics of flawed respectively accurate types. The same expressions show that, even in informational-entropic approach, the measuring errors are higher in cases with flawed characteristics comparatively with the ones having accurate features.

Now let us note some things about the comprehensive description of measuring errors in cases approached in Subsection 5.4 and in Appendix G, regarding of quantum observables with continuous spectra. The corresponding measurements, depicted through the transformations (40), can be associated with the following informational entopies

$$\begin{aligned} \mathcal{H}_\eta(\rho) &= -\int_{-\infty}^{+\infty} \rho_\eta(x) \cdot \ln\left(\rho_\eta(x)\right) \cdot dx, \\ \mathcal{H}_\eta(|j|) &= -\int_{-\infty}^{+\infty} \left|j_\eta(x)\right| \cdot \ln\left(\left|j_\eta(x)\right|\right) \cdot dx, \end{aligned} \tag{H.11}$$

where $\eta = in, out$. Related with the above entropies can be introduced the next comprehensive error indicators

$$\begin{aligned} \mathcal{E}\{H(\rho)\} &= \mathcal{H}_{out}(\rho) - \mathcal{H}_{in}(\rho) \ , \\ \mathcal{E}\{\mathcal{H}(|j|)\} &= \mathcal{H}_{out}(|j|) - \mathcal{H}_{in}(|j|)\,. \end{aligned} \tag{H.12}$$

Through some simple calculations (completely similar to the ones used in (H.2) and (H.3)) one finds that the error indicators (H.11) satisfy the relations

$$\mathcal{E}\{\mathcal{H}(\rho)\} \geqslant 0, \quad \mathcal{E}\{\mathcal{H}(|j|)\}\,. \geqslant 0 \tag{H.13}$$

These relations with equalities or inequalities refer to the cases of ideal respectively non-ideal measurements.

In particular case of measurement illustrated in Appendix G, associated with the doubly stochastic kernels (G.2) and (G.3), the error indicators (H.12) become

$$\begin{aligned} \mathcal{E}\{\mathcal{H}(\rho)\} &= \ln\sqrt{\frac{\sigma^2 + \gamma^2}{\sigma^2}} \ \approx \ \frac{1}{2}\left(\frac{\gamma}{\sigma}\right)^2, \\ \mathcal{E}\{\mathcal{H}(|j|)\} &= \ln\sqrt{\frac{\sigma^2 + \lambda^2}{\sigma^2}} \ \approx \ \frac{1}{2}\left(\frac{\lambda}{\sigma}\right)^2. \end{aligned} \tag{H.14}$$

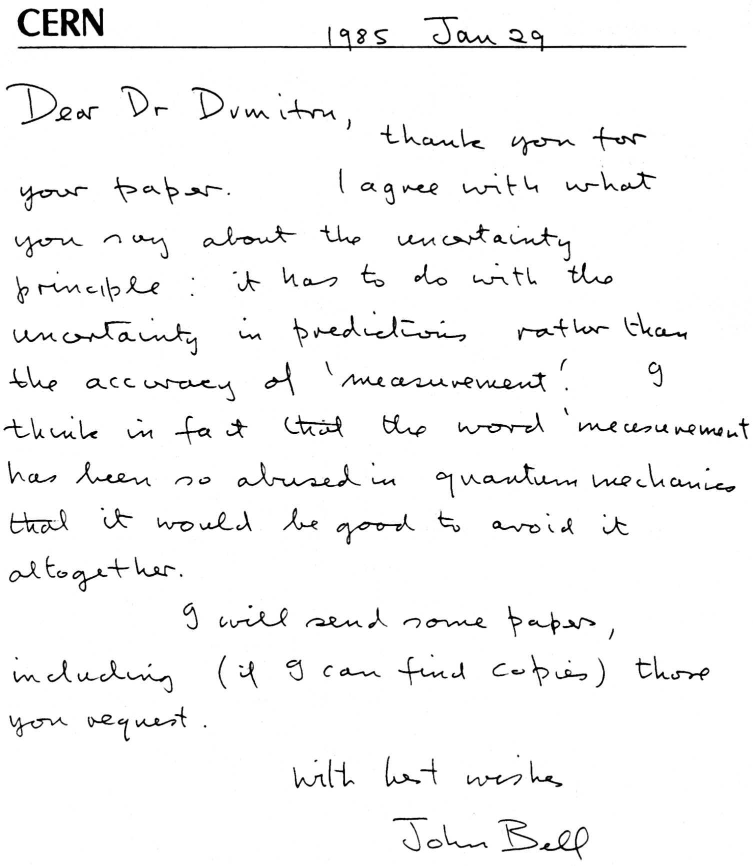
The last expressions of these indicators imply the approximations $\gamma \ll \sigma$ and $\lambda \ll \sigma$, specific to the supposition that measuring devices have high accuracies. Of course that the cases with $\gamma = 0$ and $\lambda = 0$ depict the ideal measurements.

In the case of a harmonic oscillator, mentioned in the end of Appendix G, the first error indicator from (H.12) get the expression

$$\mathcal{E}\{\mathcal{H}(\rho)\} = \ln\sqrt{\frac{\hbar + 2m\omega\gamma^2}{\hbar}} \ \approx \ \frac{m\omega}{\hbar}\,\gamma^2. \tag{H.15}$$

Submitted on November 26, 2020

I: A private letter from the late scientist J. S. Bell to the present author



CERN

1985 Jan 29

Dear Dr Dumitru, thank you for your paper. I agree with what you say about the uncertainty principle: it has to do with the uncertainty in predictions rather than the accuracy of 'measurement'. I think in fact ~~that~~ the word 'measurement has been so abused in quantum mechanics ~~that~~ it would be good to avoid it altogether.

I will send some papers, including (if I can find copies) those you request.

With best wishes

John Bell

References

1. De Witt B.S., Graham N. Resource letter IQM-1 – on the interpretation of quantum mechanics. *Am. J. Phys.*, 1971, v.39, 724–738 .
2. Nilson D.R. Bibliography on the history and philosophy of Quantum Mechanics. In: Suppes P. (ed.) Logic and probability in quantum mechanics, D. Reidel, Dordrecht, 1976.
3. Balentine L.E. Resource letter IQM-2 — Foundations of quantum mechanics since the Bell inequalities. *Am. J. Phys.*, 1987, v.55, 785–792.
4. Dodonov V.V., Man'ko V.I. Generalization of uncertainty relations in quantum mechanics. *Proc. Lebedeev Phys. Institute*, 1987, v.183, 5–70. English version, in: Markov M.A. (ed.) Invariants and evolution of non stationary quantum systems. Nova Science, New York, 1989, 3–101.
5. Auletta G. Foundations and interpretation of quantum mechanics. World Scientific, Singapore, 2000.
6. Cabello A. Bibliographic guide to the foundations of quantum mechanics and quantum information. arXiv: quant-ph/0012089v12.
7. Busch P., Heinonen T., Lahti P. Heisenberg's uncertainty principle. *Physics Reports*, 2007, v.452, 155–176.
8. Sen D., The uncertainty relations in quantum mechanics. *Current Science*, 2014, v.10, 203–218.
9. Li J.-L., Qiao C.-F. Reformulating the quantum uncertainty relation. *Sci. Rep.*, 2015 v.5, 12708.
10. Hilgevoord J., Uffink J. The uncertainty principle. In: Zalta E.N. (ed.) The Stanford Encyclopedia of Philosophy, Winter 2016 Edition.
11. Jijnasu V. The uncertainty principle — a simplified review of the four versions. *Studies in History and Philosophy of Modern Physics*, 2016, v.55, 62–71.
12. Werner R.F., Farrelly T. Uncertainty from Heisenberg to Today. arXiv: 1904.06139v1.
13. Martens H. Uncertainty Principle (PhD thesis). Eindhoven, Technical University, 1991.
14. Bunge M. The interpretation of Heisenberg inequalities. In: Pfepfer H. (ed.) Denken und Umdenken-zu Werk und Werkung von Werner Heisenberg, R. Piper, Munchen, 1977.
15. Philosophy. Cambridge Dictionary.
16. Kupczynski M. Invited Comment: Quantum mechanics and modelling of physical reality. *Phys. Scr.*, 2018, v.93, 123001, arXiv: 1804.02288v2.
17. Dumitru S. Uncertainty relations or correlation relations? *Epistemological Letters*, 1977, v.15, 1–78.
18. Dumitru S. Fluctuations but not uncertainties — deconspiration of some confusions. In: Datta B., Dutta M. (eds.) Recent Advances in Statistical Physics (Proc. Int. Bose Symposium on Statistical Physics, Dec. 28–31, 1984, Calcutta, India), World Scientific, Singapore, 1987, 122–151.
19. Dumitru S. L_z–φ uncertainty relation versus torsion pendulum example and the failure of a vision. *Rev. Roum. Phys.*, 1988, v.33, 11–45.
20. Dumitru S. Reconsideration of the uncertainty relations and quantum measurements. *Progress in Physics*, 2008, v.4, issue 2, 50–68; arXiv: 1205.3892v1.
21. Dumitru S. Do the uncertainty relations really have crucial significances for physics? *Progress in Physics*, 2010, v.6, issue 4, 25–29; arXiv: 1005.0381.
22. Occam's razor. Encyclopedia Britannica.
23. Dirac P.A.M. The evolution of the physicist's picture of Nature. *Scientific American*, 1963, v.208, 45–53.
24. Precept. The Free Dictionary.
25. Heisenberg W. Uber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Phys.*, 1927, v.43, 172–198; English translations: (i) Heisenberg W. The physical content of quantum kinematics and mechanics. In: Wheeler J.A., Zurek W.H. (eds.) Quantum Theory and Measurements. Princeton University Press, Princeton, 1983, 62–86, or (ii) Heisenberg W. The Actual content Of quantum theoretical kinematics and mechanics. Nasa Technical Memorandum, NASA TM 77379.
26. Heisenberg W. The physical principles of quantum theory. First German Edition, Leipzig, 1930; English version, Dover Publ., New York, 1949.
27. Robertson H.P. The uncertainty principle. *Phys. Rev.*, 1929, v.34, 163–164.
28. Schrödinger E. About Heisenberg uncertainty relation. *Proceedings of the Prussian Academy of Sciences, Physics-Mathematical Section*, 1930, v.19, 296–303; English versions In: (i) *Bulg. J. Phys.*, 1999, v.26, issue 5/6, 193–203; (ii) arXiv: quantph/9903100 (annotation by A. Angelow and M.-C. Batoni).
29. Bransden B.H., Joachain C.J. Quantum Mechanics. 2nd edition, Pearson Educational Limited, Essex (England), 2000.
30. Griffiths D.J. Introduction to Quantum Mechanics. Second Edition, Pearson Education Inc., 2005.
31. Schwabl F. Quantum Mechanics. Fourth Edition, Springer-Verlag, Berlin, 2007.
32. Dumitru S. Quantum Mechanics — Consistent Probabilistic Formulation. Matrix Rom, Bucharest, 2006 (*in Romanian*).
33. Born M., Wolf T. Principles of Optics. Pergamon Press, Oxford, 1986.
34. Ryachodhuri C. Heisenberg's microscope a misleading illustration. *Found. Phys.*, 1978, v.8, 845–849 .
35. Scheer J. et al. A possible evidence for a violation of Heisenberg's position-momentum uncertainty relation. *Found. Phys. Lett.*, 1989, v.2, 71–79.
36. Super-resolution microscopy. Wikipedia.
37. Hemmer P., Ben-Benjamin J.S. Invited Comment: The quest for ultimate super resolution. *Phys. Scr.*, 2016, v.91, 093003.
38. Zhou S., Jiang L. A modern description of Rayleigh's criterion. arXiv: 1801.02917v2.
39. Hofer W.A. Heisenberg uncertainty and the scanning tunneling microscope. *Frontiers of Physics*, 2012, v.7, 218–222; arXiv: 1105.3914v3.
40. Rozema L.A., Darabi A., Mahler D.H., Hayat A., Soudagar Y., Steinberg A.M. Violation of Heisenberg's measurement-disturbance relationship by weak measurements. *Phys. Rev. Lett.*, 2012, v.109, 100404; arXiv: 1208.0034v2.
41. Popper K. Quantum mechanics without the "observer". In: Bunge M. (ed.) Quantum Theory and Reality, Springer, Berlin, 1967, 7–44.
42. Judge D. On the uncertainty relation for L_z and φ. *Phys. Lett.*, 1963, v.5, 189.
43. Carruthers P., Nieto M.M. Phase and angle variables in quantum mechanics. *Rev. Mod. Phys.*, 1968, v.40, 411–440.
44. Roy C.L., Sannigrahi A.B. Uncertainty relation between angular momentum and angle variable. *Am. J. Phys.*, 1979, v.47, 965.
45. Galitski V., Karnakov B., Kogan V.. Problemes de mecanique quantique. Mir, Moscou, 1985.
46. Barnett S.M., Pegg D.T. Quantum theory of rotation angles. *Phys. Rev. A.*, 1990, v.41, issue 7, 3427–3435.
47. Dumitru S. Compatibility versus commutativity — the intriguing case of angular momentum — azimuthal angle. In: Dodonov V.V., Man'ko V.I., (eds.) *Proc. Lebedev Phys. Inst. Acad. Sci. USSR*, v.187; Quantum field theory, quantum mechanics and quantum optics. Pt.1. Symmetries and algebraic structures in physics. Nova Sci. Pub., New York, 1991, 243–246.
48. Nieto M.M. Quantum phase and quantum phase operators: some physics and some history. *Phys. Scr.*, 1993, v.T48, 5–12; arXiv: hep-th/9304036.

49. Lynch R. The quantum phase problem — a critical review. *Physics Reports*, 1995, v.256, 367–436.
50. Kovalski K., Rebielinski J. On the uncertainty relations and squeezed states for the quantum mechanics on a circle. *J. Phys. A: Math. Gen.*, 2002, v.35, 1405–1414.
51. Trifonov D.A. On the position uncertainty measure on the circle. *J. Phys. A: Math. Gen.*, 2003, v.36, 11873–11879.
52. Franke-Arnold S., Barnet S.M., Yao E., Leach J., Courtial J., Padgett M. *New Journal of Physics*, 2004, v.6, issue 103, 1–8.
53. Pegg D.T., Barnett S.M., Zambrini R., Franke-Arnold S., Padgett M. Minimum uncertainty states of angular momentum and angular position. *New Journal of Physics*, 2005, v.7, issue 62, 1–20.
54. Kastrup H.A. Quantization of the canonically conjugate pair angle and orbital angular momentum. *Phys. Rev. A*, 2006, v.73, 052104; arXiv: quant-ph/0510234.
55. Dumitru S. A possible general approach regarding the conformability of angular observables with mathematical rules of quantum mechanics. *Progress in Physics*, 2008, v.4, issue 1, 25–30; arXiv: quant-ph/0602147.
56. Doubrovine B., Novikov S., Fomenko A., Geometrie contemporaine. Premier Partie, Mir, Moscou, 1982.
57. Molecular vibration. Wikipedia.
58. Busch P., Lahti P.J. The complementarity of quantum observables: theory and experiments. *Rivista del Nuovo Cimento*, 1995, v.18, issue 4, 1–27; arXiv: quant-ph/0406132.
59. Sharatchandra H. S. Phase of the quantum oscillator. arXiv: quant-ph/9710020.
60. Busch P., Lahti P., Pellonpaa J.-P., Ylinen K. Are number and phase complementary observables? *Journal of Physics A: Math. Gen.*, 2001, v.34, 5923–5935; arXiv: quant-ph/0105036.
61. Kitajima S., Shingu-Yano M., Shibata F. Number-phase uncertainty relation. *J. Phys. Soc. Japan*, 2003, v.72, 2133–2136.
62. Bauer M., Mello P.A. The time-energy uncertainty relation. *Ann. Phys.*, 1978, v.111, 38–60.
63. Bush P. The time energy — uncertainty relation. In: Muga J.G., Sala M.R., Egusquiza I.L. (eds.) Time in Quantum Mechanics, Springer, Berlin, 2002, 69–98.
64. Urbanowski K. Remarks on the time — energy uncertainty relation. arXiv: 1810.11462v1.
65. Angular resolution. Wikipedia.
66. Folland G.B., Sitaram A. The uncertainty principle: a mathematical survey. *The Journal of Fourier Analysis and Applications*, 1997, v.3, issue 3, 207–238.
67. Hall M. What is the Gabor uncertainty principle? https://agilescientific.com/blog/2014/1/15/what-is-the-gabor-uncertainty-principle.html
68. Munster A. Statistical Thermodynamics. Springer, Berlin, 1965.
69. Landau L., Lifchitz E. Physique Statistique. Premiere partie, Mir, Moscow 1984.
70. Dumitru S. Fluctuations and thermodynamic inequalities. *Phys. Scr.*, 1974, v.10, 101–103.
71. Dumitru S., Boer A. Fluctuations in the presence of fields — phenomenological Gaussian approximation and a class of thermodynamic inequalities. *Phys. Rev. E*, 2001, v.64, 021108; arXiv: cond-mat/0011444v1.
72. Boer A., Dumitru S. Higher-order correlations for fluctuations in the presence of fields. *Phys. Rev. E*, 2002, v.66, 046116; arXiv: cond-mat/0201171v2.
73. Dumitru S. The Plank and Boltzmann constants as similar generic indicators of stochasticity — some conceptual implications of quantum-nonquantum analogies. *Physics Essays*, 1993, v.6, 5–20; CERN Central Library PRE 32387 rev.; This article was awarded in 1995 with the Dragomir Hurmuzescu prize of the Romanian Academy.
74. Jancel R. Foundations of Classical and Quantum Statistical Mechanics. Pergamon Press, New York, 1973.
75. Density matrix. Wikipedia.
76. Zubarev D. N. Nonequilibrium statistical thermodynamics. Nauka, Moscow, 1974 (*in Russian*); English version: Consultants Bureau, New York, 1974.
77. De Groot S.R., Mazur P. Nonequilibrium Thermodynamics. North-Holland, Amsterdam, 1962.
78. Uffink J., Van Lith J. Thermodynamic uncertainty relations. *Found. Phys.*, 1999, v.29, 655–692; arXiv: cond-mat/9806102v1.
79. Artamonov A.A., Plotnikov E.V. Thermodynamic uncertainty relation as a fundamental aspect of quantum thermodynamics. *Bulletin of the Tomsk Polytechnic University — Resource-Efficient Technologies*, 2018 v.1, 17–29.
80. Tyablikov S.V. Metods of Quantum Theory of Magnetism. Nauka, Moscow, 1975 (*in Russian*).
81. Schwabl F. Statistical Mechanics. Springer, Berlin, 2002.
82. Observer effect (physics). Wikipedia.
83. Bell J.S. A private letter from the late scientist J. S. Bell to the present author (dated January 29, 1985). See a copy in Appendix I; also in: arxiv:quant-ph/0004013v1.
84. Bell J.S. Against "measurement". *Physics World*, 1990, v.3, 33–40. Article reprinted also in some books containing Bell's writings.
85. Category: Principles. Wikipedia.
86. Busch P., Lahti P., Mittelstaedt P. The Quantum Theory of Measurement. Second Edition, Springer, Berlin, 1996.
87. Svensson B.E.Y. Pedagogical review of quantum measurement theory with an emphasis on weak measurements. *Quanta*, 2013, v.2, issue 1, 18–49.
88. Wayne M. Philosophical issues in quantum theory. In: Zalta E.N. (ed.) The Stanford Encyclopedia of Philosophy, Spring 2017 Edition.
89. Hossenfelder S. At the frontier of knowledge. arXiv: 1001.3538v1.
90. Dumitru S., Verriest E.I. Behaviour patterns of observables in quantum-classical limit. *Int. J. Theor. Physics*, 1995, v.34, issue 8, 1785–1790.
91. Dumitru S., Boer A. On the measurements regarding random observables, *Rom. Journal. Phys.*, 2008, v.53, issue 9/10, 1111–1116.
92. Dumitru S. Routes of quantum mechanics theories. *Progress in Physics*, 2012, v.8, issue 3, L1.
93. Dumitru S. Caducity of idea about wave function collapse as well new views on Schrödinger's cat and quantum measurements. *Progress in Physics*, 2013, v.9, issue 1, 63–68; arXiv: 1210.4121v3.
94. Dumitru S. Views about the "Oxford questions", wave function collapse and Schrödinger's cat — are they real scientific topics or plain fictions? *Progress in Physics*, 2014, v.10, issue 2, 111–113; arXiv: 1311.2581.
95. Klyshko D.N., Lipkine A.I. About the "reduction of wave function", quantum theory of measurement, and "incomprehension" of quantum mechanics. *Electronic Journal "Investigated in Russia"*, 2000, 703–735.
96. Albertson J. Quantum-mechanical measurement operator. *Phys. Rev.*, 1963, v.129, 940–943.
97. Communication channel. Wikipedia.
98. Doubly stochastic matrix. Wikipedia.
99. Stamatescu I.-O. Wave function collapse. In: Greenberger D., Hentschel K., Weinert F. (eds.) Compendium of Quantum Physics, Concepts, Experiments, History and Philosophy. Springer, Dordrecht, 2009, 813–822.
100. Omnes R. Decoherence and wave function collapse. *Found. Phys.*, 2011, v.41, 1857–1880.
101. Weinberg S. Collapse of the state vector. *Phys. Rev. A*, 2012, v.85, 062116; arXiv: 1109.6462.

102. Bassi A., Lochan K., Satin S., Singh T.P., Ulbricht H. Models of wave-function collapse, underlying theories, and experimental tests. *Rev. Mod. Phys.*, 2013, v.85, 471–527; arXiv: 1204.4325.
103. Ghirardi G. Collapse theories. In: Zalta E.N. (ed.) The Stanford Encyclopedia of Philosophy, Spring 2016 Edition.
104. Schrödinger E. The present situation in quantum mechanics: A translation of Schrödinger's "cat paradox paper" (1935), Translator: Trimmer J.D. *Proc. Am. Philos. Soc*, 1980, v.124, 323–338.
105. Hobson A. Review and suggested resolution of the problem of Schrödinger's cat. *Contemporary Physics*, 2018, v.59, issue 1, 16–30; arXiv: 1711.11082v1.
106. Schrödinger's cat. Wikipedia.
107. Dumitru S. Phenomenological theory of recorded fluctuations. *Phys. Lett. A.*, 1974, v.48, 109–110.
108. Boyarsky A., Gora P. Laws of chaos — invariant measures and dynamical systems in one dimension. Springer Science — Business Media, 2012, 240.
109. Circular error probable. Wikipedia.
110. The Schrödinger's cat thought experiment. http://www.markstokoe.ca/files/Dfp_Mono_v6_5.pdf
111. Dumitru S. Are the higher order correlations resistant against additional noises? *Optik*, 1999, v.110, 110–112.
112. Einstein A., Podolsky B., Rosen N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 1935, v.47, 777–780.
113. Cantrell C.D., Scully M. The EPR paradox revisited. *Physics Reports*, 1978, v.43, issue 13, 499–508.
114. Schlosshauer M. Prologue to: Elegance and Enigma — The Quantum Interviews. Springer, Berlin, 2011, 19.
115. Ozawa M. Heisenberg's uncertainty relation — violation and reformulation, *Journal of Physics: Conference Series*, 2014, v.504, 012024.
116. Okamura K., Ozawa M. Universally valid uncertainty relations in general quantum systems. arXiv: 1808.10615v2.
117. Huang X., Zhang T., Jing N. On uncertainty relations in the product form. arXiv: quant-ph/2003.10696v1.
118. Song Q.-C. et al. A stronger multi-observable uncertainty relation. *Sci. Rep.*, 2017, v.7, 44764.
119. Homayouni S. Some Generalizations of the Heisenberg Uncertainty Principle (PhD thesis). Ottawa, Ontario, Canada, Carleton University, 2011.
120. Nguyen Q.H., Bui Q.T. Mathematical uncertainty relations and their generalization for multiple incompatible observables. *VNU Journal of Science: Mathematics Physics*, 2017, v.33, 29–34.
121. Madelung E. Die Mathematischen Hilfsmittel der Physikers. Springer Verlag, Berlin, 1957; Russian version: Gosizdat, Moscow, 1961.

Laser Action in the Stellar Atmospheres of Wolf-Rayet Stars and Quasi-Stellar Objects (QSOs)

Pierre A. Millette

E-mail: pierre.millette@uottawa.ca, Ottawa, Canada

In this paper, we reconsider the little-known but critically important physical process of laser action occurring in the stellar atmospheres of Wolf-Rayet stars and, by extension, of QSOs, also known as quasars in the cosmological context. We review the use of the Collisional-Radiative (non-LTE) model for hydrogenic and lithium-like ions to calculate the energy level populations and the existing results for He I, He II, C III and C IV, and for N V and O VI. We review the details of laser action in Wolf-Rayet stars, as well as in QSOs. We note that taking QSOs to be local stellar objects eliminates the problems associated with their cosmological interpretation. We propose that the terminology *quasar* be used to refer to the cosmological interpretation and *QSO* to refer to the stellar interpretation of Quasi-Stellar Objects. We introduce a new star type Q for QSOs, similar to the star type W for Wolf-Rayet stars. We expand the Hertzsprung-Russell diagram to include more massive and hotter stars of type Q and W beyond the stars of type O B. The main sequence thus starts with stars of type Q W O B, followed by the rest of the main sequence. Finally, we note the effort that will be required to understand the classification and evolution of stars of type Q, as has been achieved for Wolf-Rayet stars.

1 Introduction

In this paper, we reconsider a little-known but critically important physical process occurring in the stellar atmospheres of Wolf-Rayet stars and, by extension, of Quasi-Stellar Objects (QSOs), also known as quasars in the cosmological context. Wolf-Rayet stars are known to have an expanding envelope of hot ionized gases, as the stellar atmosphere of the star expands, resulting in mass loss.

If the speed of expansion is low, the expansion will be closer to being isothermal, but as the speed of expansion increases, the expansion will become adiabatic. Under those conditions, as the plasma cools, population inversions will occur in the ionic energy levels due to free electron-ion recombination in higher ionic excited states. Some ionic energy level transitions will undergo laser action [1] resulting in spectra dominated by a small number of strong broad emission lines, which becomes even more evident in QSOs.

2 Wolf-Rayet stars

Wolf-Rayet stars [2] are a type of stars that, like the supergiants, have extended atmospheres whose thickness is an appreciable fraction of their stellar radius [3, p. 243]. Characteristic features in the visible spectra of many O and early B stars, particularly supergiants, and WR stars provide evidence that these objects have extensive envelopes, and that the material generating the lines is flowing outward from the stellar photosphere.

The number of WR stars in our galaxy is small: the 2001 VIIth catalog of galactic WR stars gave the number at 227 stars, comprised of 127 WN stars, 87 WC stars, 10 WN/WC stars and 3 WO stars [4]. The subtypes are covered in the spectra discussion later in this section. A 2006 update added another 72 WR stars, including 45 WN stars, 26 WC stars and one WO star [5]. The latest number from the August 2020 Galactic Wolf Rayet Catalogue v1.25 is 667 WR stars [6].

The existence of large-scale, rapid, and sometimes violent expansions of stellar atmospheres is well-established observationally [3, p. 471]. Beals [7, 8] first recognized that the great breadths of lines in WR spectra, indicating velocities of the order of 3 000 km/s, could be interpreted in terms of rapid outflow of material. His suggestion that the flow was driven by radiation pressure is supported by current dynamical models. Further evidence for mass loss is provided by infrared and radio continuum observations of several OB and WR stars, which are most readily interpreted in terms of free-free emission from an extended, optically-thick envelope having a density profile consistent with steady outflow of the stellar atmosphere [3, p. 550–551].

We know today, from a variety of observational evidence from spacecraft and ground-based observatories, that in the WR and Of stars and in many early-type supergiants, there are massive trans-sonic stellar winds, that have very small outward velocities in the deeper layers of the stars, but a large outward acceleration producing very large velocities ($v/c \approx 0.01$) at great distances from the stars [3, pp. 471–472,550]. These flows are driven by radiation pressure acting on the stellar atmosphere [3, p. 523].

Mass loss in stellar winds, particularly in the early-type OB supergiants and WR stars, is well established [3, pp. 266, 523]. The analysis of line profiles and infrared emissions imply estimated mass loss rates $\mathcal{M}$ of order 10^{-6} to 10^{-5} $\mathcal{M}_{\odot}$/year for O stars and perhaps up to 10^{-4} $\mathcal{M}_{\odot}$/year for

WR stars [9, p. 628]. For comparison, mass loss rates for the solar wind is about 10^{-14} $\mathcal{M}_\odot$/year. The flow velocities rise from close to zero in the stellar photosphere to highly supersonic values within one stellar radius from the surface. The 3 000 km/s flow is thought to be driven by momentum input to the ionized gas from the intense radiation force exerted by the strong spectrum lines of these extremely luminous stars.

Series of extremely strong emission lines can be observed in the spectra of WR stars. The spectra fall into two broad classes: WN, which have prominent lines of nitrogen N and helium He ions, with a very strong He II Pickering series ($n = 4 \rightarrow n'$), and essentially no lines of carbon C; and WC, where the lines of carbon C and oxygen O are prominent along with the helium He ions, while those of nitrogen N seem to be practically absent [3, p. 485]. An additional subtype WO with strong O VI lines has also recently been added as a separate subtype. The spectra are characterized by the dominance of emission lines, notable for the almost total absence of hydrogen H lines [10].

3 Laser action in stellar atmospheres

In initial modelling calculations, Castor [11] used the escape probability method of basic Sobolev theory to treat the transfer of line radiation in a stellar envelope to provide a coarse analysis of the spectral line formation in Wolf-Rayet stars for a line formed in a two-level atom [3, p. 471–472]. He then used this analysis to calculate the populations of the first thirty levels of hydrogen-like He II ions under statistical equilibrium with all radiative and collisional transitions included [12]. He also applied this analysis to 14 terms and all allowed transitions of helium-like C III ions; no case of laser action was found in the calculations as the existing atomic processes used did not provide sufficient pumping of the excited levels to maintain population inversion [13].

Mihalas [3, p. 485–490] carries out a complete multilevel analysis of the spectrum of an ion using statistical equilibrium equations that consider the radiative and collisional processes contributing to the population of each ionic level under consideration. Typically, the only free parameters in this analysis are T_e, the temperature of the free electrons corresponding to the envelope temperature, n_e, the free electron number density and n_{atom}, the total number density of the species (element) under consideration. The analysis is done under Local Thermodynamic Equilibrium (LTE) conditions, that is under conditions in which each volume element of the plasma fulfills all thermodynamic equilibrium laws derived for plasmas in complete thermodynamic equilibrium (CTE) except for Planck's radiation law [40, p. 12–13].

3.1 Plasma lasers

The possibility of using a recombining plasma as an amplifying medium of electromagnetic radiation was first suggested by Gudzenko and Shelepin [14]. Calculations performed on a hydrogen plasma [15, 17] subsequently confirmed this suggestion. Such plasmas are called plasma lasers [18].

We consider the basic principles of operation of a plasma laser. The mean time between electron collisions determines the rate of establishment of the electron temperature within a plasma. The smallness of the time between elastic collisions in a dense plasma thus makes it possible, in principle, to rapidly reduce the electron temperature of such a plasma. For example, in plasma densities of order $n_i \sim n_e \sim 10^{15} - 10^{16}\,\mathrm{cm}^{-3}$, a single distribution of the electrons is established in a time of order $\tau \sim 10^{-11} - 10^{-10}$ s [14], where n_i is the ion number density.

Rapid cooling of a strongly ionized plasma results in rapid recombination of the electrons and the ions into highly excited ionic states. The subsequent relaxation of the electrons to the ground state by spontaneous and non-radiative transitions occurs in a time which, for the estimated values of the plasma parameters used in this work, is larger than 10^{-7} s. At those densities, electron-ion recombination occurs by three-body recombination in a time shorter than 10^{-7} s such that a rapid filling-up of the upper excited levels of the ions occurs. Furthermore, since recombination into highly excited states occurs much more rapidly than into lower states, the establishment of large population inversions is favored.

When large population inversions have been established in the excited levels, the plasma is said to be in a stationary drainage state. It is still substantially ionized. As an example of the time involved, Gudzenko et al. [15] find that for a dense low temperature plasma ($T_e \sim 1000 - 6000$ K and n_e-bound and free states $\sim 10^{13} - 10^{16}\,\mathrm{cm}^{-3}$), cooled by a factor of twenty, stationary drainage of the excited discrete levels is established in a time $\sim 10^{-8} - 10^{-7}$ s. Stationary drainage is maintained for a time $\sim 10^{-5}$ s, and is followed by a stage in which the plasma is weakly ionized and the population densities of its levels return to normal. Gudzenko et al. [17] find that the above conditions can be significantly relaxed; for example, the cooling can be done more slowly or by stages [40, p. 42–43].

3.2 Adiabatic cooling of a plasma

Various mechanisms of free electron cooling can be used. The method of interest to us, rapid cooling of a plasma by adiabatic expansion, was first investigated by Gudzenko et al. [16] both for magnetized and unmagnetized plasmas.

An example of this cooling mechanism is the adiabatic expansion of a plasma jet in a vacuum. The advantage of this method is that continuous amplification, and thus continuous operation of a laser is possible due to the fact that the different stages of the recombining plasma decay at different times. Thus, as the plasma expands, the stages of the recombination process outlined in the previous Section §3.1 are spread over space and the de-excited medium is thus removed from the active lasing zone. Experimental evidence of laser action due to the adiabatic expansion of highly ionized hydrogen or hydrogenic plasmas has been given by, for example, [19] and [20].

Under adiabatic expansion conditions, the density n and the temperature T of a gas are related by [17]

$$T\, n^{1-\gamma} = \text{constant} \tag{1}$$

where

$$\gamma = c_P / c_V \tag{2}$$

is the ratio of the specific heat at constant pressure c_P and the specific heat at constant volume c_V. For a monatomic gas and for a fully ionized plasma of hydrogen, we use [17]

$$\gamma = 5/3\,. \tag{3}$$

However, it should be noted that the actual value of γ for a plasma is slightly smaller than 5/3. Denoting the initial density and temperature of the plasma by n_0 and T_0 respectively, and the final density and temperature by n and T respectively, we characterize the expansion by the factor

$$f_E = \frac{n_0}{n} > 1 \tag{4}$$

and the ensuing cooling of the plasma by the factor

$$f_C = \frac{T_0}{T} > 1\,. \tag{5}$$

Then from (1), we have the relation

$$f_C = f_E^{\gamma-1} \tag{6}$$

under adiabatic expansion conditions. In this work, we use $f_c = 5$; then from (6) and (3), $f_E = 11.2$ [40, p. 43–44].

4 The Collisional-Radiative (non-LTE) model

To calculate the non-equilibrium population of the ionic energy levels, we need to use a model that applies to non-LTE plasmas. The Collisional-Radiative (CR) non-LTE model was first proposed and applied to hydrogenic ions by Bates et al. [21,22] and subsequently used by Bates and Kingston [23] and McWhirter and Hearn [24]. It was first applied to helium by Drawin and Emard [25], to lithium by Gordiets et al. [26], and to cesium by Norcross and Stone [27].

The population densities of the energy levels of ions in non-LTE plasmas must be obtained from the rate coefficients of the individual collisional and radiative processes occurring within the plasma, as summarized in Fig. 1. The physical processes included in the CR model include:

- Collisional ionization by electron impact
  Rate coefficient: $S_p(T)\,\text{cm}^3\text{s}^{-1}$
  Number of processes: $n_p\, n_e\, S_p(T)\,\text{cm}^{-3}\text{s}^{-1}$
- Three-body recombination
  Rate coefficient: $\alpha_p(T)\,\text{cm}^6\text{s}^{-1}$
  Number of processes: $n_e^2\, n_i\, \alpha_p(T)\,\text{cm}^{-3}\text{s}^{-1}$

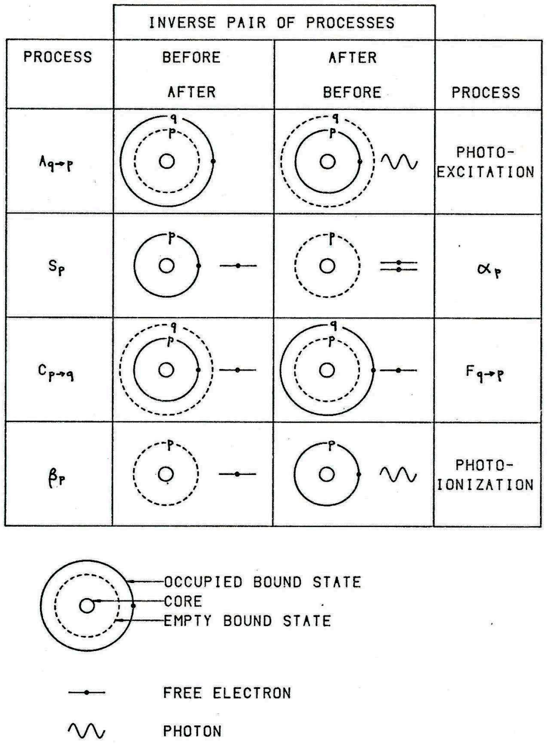


FIGURE 1.2 - PHYSICAL PROCESSES OCCURRING IN A PLASMA.

Fig. 1: This figure provides a summary of the collisional and radiative processes occurring within the plasma, where p and q are ionic energy state labels; $p \gtrless q$ [40, p. 21].

- Radiative recombination
  Rate coefficient: $\beta_p(T)\,\text{cm}^3\text{s}^{-1}$
  Number of processes: $n_e\, n_i\, \beta_p(T)\,\text{cm}^{-3}\text{s}^{-1}$
- Collisional excitation by electron impact ($p < q$)
  Rate coefficient: $C_{p\to q}(T)\,\text{cm}^3\text{s}^{-1}$
  Number of processes: $n_p\, n_e\, C_{p\to q}(T)\,\text{cm}^{-3}\text{s}^{-1}$
- Collisional de-excitation by electron impact ($p < q$)
  Rate coefficient: $F_{q\to p}(T)\,\text{cm}^3\text{s}^{-1}$
  Number of processes: $n_q\, n_e\, F_{q\to p}(T)\,\text{cm}^{-3}\text{s}^{-1}$
- Spontaneous transition ($p < q$)
  Rate coefficient: $A_{q\to p}\,\text{s}^{-1}$
  Number of processes: $n_q\, A_{q\to p}\,\text{cm}^{-3}\text{s}^{-1}$

The plasma is assumed to be optically thin such that all radiation emitted within the plasma escapes without being absorbed. The following physical processes are thus neglected:

- Photoexcitation ($p < q$),
- Photoionization.

The differential equation describing the time variation of the population density of a given ionic level p is then given by

$$\begin{aligned} \frac{dn_p}{dt} &= (\text{electrons entering level } p) - \\ &- (\text{electrons leaving level } p)\,. \end{aligned} \tag{7}$$

The terms of (7) in parentheses include contributions from all levels $q < p$, $q > p$, and continuum states. Substituting for the collisional and radiative processes considered above, we obtain the differential equation

$$\begin{aligned} \dot{n}_p &= \sum_{q=1}^{p-1} C_{q\to p}\, n_e\, n_q - \\ &-\left[\left(\sum_{q=1}^{p-1} F_{p\to q} + S_p + \sum_{q=p+1}^{\infty} C_{p\to q}\right) n_e + \sum_{q=1}^{p-1} A_{p\to q}\right] n_p + \\ &+ \sum_{q=p+1}^{\infty}\left(F_{q\to p}\, n_e + A_{q\to p}\right) n_q + \\ &+ \left(\alpha_p\, n_e + \beta_p\right) n_e\, n_i \end{aligned} \tag{8}$$

where the dot over n_p represents differentiation with respect to time. There is such an equation for each and every level $p = 1, 2, ..., \infty$ of the ion. We thus obtain an infinite number of coupled first order differential equations in the population densities of the discrete levels of the ion.

The population density of level p, n_p, is normalized with the Saha equilibrium population density of level p, n_p^E, [28, p. 154] [29, p. 135]

$$\rho_p = \frac{n_p}{n_p^E}\,, \tag{9}$$

with n_p^E given by

$$n_p^E = Z_p(T)\, n_i\, n_e\,, \tag{10}$$

where

$$Z_p(T) = \frac{\omega_p}{u_i}\,\frac{h^3}{2(2\pi mkT)^{3/2}}\, e^{I_p/kT}\,, \tag{11}$$

ω_p is the statistical weight of level p, u_i is the ionic partition function, and I_p is the ionization potential of state p. For hydrogenic ions, u_i is the partition function of the bare nucleus and is given by $u_i \simeq 1$. The same holds for lithium-like ions since u_i is then the partition function of a closed shell ion.

The relative population densities of various stages of ionization n_i of a monatomic non-LTE plasma under statistical equilibrium are calculated approximately with the model of House [30]. Even though the calculations are highly simplified, the model provides a first approximation to the ionization equilibrium of monatomic plasmas of hydrogen to iron and a general method of obtaining a consistent set of relative population densities for the ionization stages of these elements.

Given that there exists a high-lying quantum state r above which the discrete levels are in LTE, the normalization (9) allows us to set the population density of these levels to be given by $\rho_{p>r} = 1$. The infinite set of equations (8) thus becomes a finite set of r coupled equations which can be solved for ρ_p, $p = 1, 2, ..., r$. The infinite sums appearing in (8) can be cut off at a sufficiently high-lying level $s > r$ above which the rate coefficients involving these states contribute little to the infinite sums of (8). For levels in LTE, detailed balancing between the collisional excitation and de-excitation processes holds and then we can use

$$n_q^E\, F_{q\to p} = n_p^E\, C_{p\to q}\,. \tag{12}$$

The set of equations (8) then becomes

$$\begin{aligned} \dot{\rho}_p &= \sum_{q=1}^{p-1} F_{p\to q}\, n_e\, \rho_q - \\ &-\left[\left(\sum_{q=1}^{p-1} F_{p\to q} + S_p + \sum_{q=p+1}^{s} C_{p\to q}\right) n_e + \sum_{q=1}^{p-1} A_{p\to q}\right] \rho_p + \\ &+ \sum_{q=p+1}^{r}\left(C_{p\to q}\, n_e + \frac{Z_q}{Z_p} A_{q\to p}\right) \rho_q + \frac{1}{Z_p}\left(\alpha_p\, n_e + \beta_p\right) + \\ &+ \sum_{q=r+1}^{s}\left(C_{p\to q}\, n_e + \frac{Z_q}{Z_p} A_{q\to p}\right);\, p = 1, 2, ..., r\,. \end{aligned} \tag{13}$$

4.1 Solution of the system of coupled differential equations

The exact solution of the system of couple differential equations (13) gives the time evolution of the population densities of the ionic levels $\rho_p(t)$, $p = 1, 2, ..., r$. This solution if of limited use. A simpler solution, known as the quasi-steady state (QSS) approximation, holds for a large class of plasmas and is used extensively in the literature (see [21, 22] and subsequent papers mentioned previously in §4). The steady state (SS) solution is obtained by putting

$$\dot{\rho}_p^{SS}(t) = 0\,;\, p = 1, 2, ..., r\,. \tag{14}$$

This time-independent solution holds when the rate at which the electrons enter level p equals the rate at which they leave level p. Once the steady state solution is established, a perturbation of the population density of level p will be followed by a return to its steady state value in a time of order

$$\tau_p \sim \left\{\left[\sum_{q=1}^{p-1} F_{p\to q} + S_p + \sum_{q=p+1}^{s} C_{p\to q}\right] n_e + \sum_{q=1}^{p-1} A_{p\to q}\right\}^{-1} \tag{15}$$

where τ_p is the relaxation time of level p.

McWhirter and Hearn [24] have calculated τ_p for a wide range of plasma parameters. They conclude that the relaxation time of the ground state is always much greater than that of any of the excited states, even if the plasma is not near its steady state. This is due to two main reasons: a) the electron collision rate coefficients between excited states are much greater than those involving the ground state; b) the ground state cannot decay by spontaneous radiative transition. Consequently, the population densities of the excited ionic levels come into equilibrium with particular values of the population densities of the ground state, of the free electrons, and of the ions in a time which is very short as compared to the ground state relaxation time. This is the basis of the QSS solution.

4.2 The population coefficients

We thus express the population densities of the excited states as a function of the ground state population density:

$$\rho_p = r_p^{(0)} + r_p^{(1)}\,\rho_1\,;\; p = 2, 3, ..., r\,. \tag{16}$$

$r_p^{(0)}$ and $r_p^{(1)}$ are called the population coefficients of level p. Furthermore, since the population densities of the excited states are in equilibrium with that of the ground state, we solve the system of coupled equations (13) by putting $\dot{\rho}_{p\geq 2} = 0$ and $\dot{\rho}_1 \neq 0$ since, in general, the ground state is not in equilibrium. In our calculations, we also assume that the free electron and ionic densities, n_e and n_i respectively, do not change substantially during the time of establishment of the QSS.

Substituting the trial solution (16) in the system of equations (13), we obtain a set of equations of the form

$$a_p + b_p\,\rho_1 = 0\,;\; p = 2, 3, ..., r\,. \tag{17}$$

The general solution of (17), for an arbitrary value of ρ_1, is $a_p = 0$ and $b_p = 0$. Before proceeding with the solution, certain limiting conditions must be imposed on the population coefficients $r_p^{(0)}$ and $r_p^{(1)}$ corresponding to the cases when $p = 1$ and $p > r$. Substituting $p = 1$ in (16), we obtain the condition $r_1^{(0)} = 0$ and $r_1^{(1)} = 1$. The other condition, which is obtained by putting $p > r$ in (16), has already been imposed on the set of equations, namely $r_{p>r}^{(0)} = 1$ and $r_{p>r}^{(1)} = 0$.

Using these conditions, we obtain the following two sets of $r - 1$ equations in the population coefficients $r_p^{(0)}$ and $r_p^{(1)}$ respectively:

$$\begin{aligned}
&\sum_{q=2}^{p-1} F_{p\to q}\, n_e\, r_q^{(0)} - \\
&-\left[\left(\sum_{q=2}^{p-1} F_{p\to q} + S_p + \sum_{q=p+1}^{s} C_{p\to q}\right) n_e + \sum_{q=1}^{p-1} A_{p\to q}\right] r_p^{(0)} + \\
&+\sum_{q=p+1}^{r}\left(C_{p\to q}\, n_e + \frac{Z_q}{Z_p} A_{q\to p}\right) r_q^{(0)} = -\frac{1}{Z_p}\left(\alpha_p\, n_e + \beta_p\right) - \\
&-\sum_{q=r+1}^{s}\left(C_{p\to q}\, n_e + \frac{Z_q}{Z_p} A_{q\to p}\right);
\end{aligned} \tag{18}$$

$$\begin{aligned}
&\sum_{q=2}^{p-1} F_{p\to q}\, n_e\, r_q^{(1)} - \\
&-\left[\left(\sum_{q=1}^{p-1} F_{p\to q} + S_p + \sum_{q=p+1}^{s} C_{p\to q}\right) n_e + \sum_{q=1}^{p-1} A_{p\to q}\right] r_p^{(1)} + \\
&+\sum_{q=p+1}^{r}\left(C_{p\to q}\, n_e + \frac{Z_q}{Z_p} A_{q\to p}\right) r_q^{(1)} = \\
&= -F_{p\to 1}\, n_e\,;\; p = 2, 3, ..., r\,.
\end{aligned} \tag{19}$$

4.3 The population densities

Once the population coefficients $r_p^{(0)}$ and $r_p^{(1)}$ have been obtained from the sets of equations (18) and (19) respectively, they are substituted in (16). For any value of ρ_1, the population densities ρ_p can then be calculated. From (9) and (10),

$$\rho_p = \frac{n_p}{Z_p\, n_i\, n_e}\,; \tag{20}$$

substituting (20) in (16), we obtain

$$n_p = Z_p\, n_i\, n_e\, r_p^{(0)} + \frac{Z_p}{Z_1}\, n_1\, r_p^{(1)}\,;\; p = 2, 3, ..., r\,. \tag{21}$$

As required by the QSS approximation, the population density of the excited state p depends on the value of the ground state population density n_1, the free electron density n_e, and the ionic density n_i. The population density per unit statistical weight is given by $y_p = n_p/\omega_p$, where ω_p is the statistical weight of level p. The population density per unit statistical weight must be used when the population densities of different states are compared.

4.4 The collisional-radiative rate coefficients

The time evolution of the population density of the ground state can be studied with (13) when $p = 1$. Substituting for ρ_p from (16), and using the previously calculated population coefficients and (9), we obtain the differential equation

$$\dot{n}_1 = -S_{CR}\, n_e\, n_1 + \alpha_{CR}\, n_e\, n_i\,. \tag{22}$$

S_{CR} and α_{CR} are called the collisional-radiative ionization and recombination rate coefficients respectively. They are the effective ionization and recombination rate coefficients of the plasma. They are related to the individual atomic rate coefficients by the following expressions:

$$\begin{aligned}
S_{CR} = S_1 + &\sum_{q=2}^{s} C_{1\to q} - \\
&-\frac{1}{Z_1\, n_e}\sum_{q=2}^{s} Z_q\left(F_{q\to 1}\, n_e + A_{q\to 1}\right) r_q^{(1)}\,;
\end{aligned} \tag{23}$$

$$\alpha_{CR} = \alpha_1\, n_e + \beta_1 + \sum_{q=2}^{s} Z_q\left(F_{q\to 1}\, n_e + A_{q\to 1}\right) r_q^{(0)}\,. \tag{24}$$

The solution of (22) can easily be shown to be given by

$$n_1(t) = \frac{\alpha_{CR}}{S_{CR}}\, n_i + \left(n_1(t=0) - \frac{\alpha_{CR}}{S_{CR}}\, n_i \right) e^{-S_{CR}\, n_e\, t} \,. \tag{25}$$

The steady state population density of the ground state is obtained in the limit $t \to \infty$:

$$n_1^{SS} = \frac{\alpha_{CR}}{S_{CR}}\, n_i \,. \tag{26}$$

4.5 Modifications for lithium-like ions

The CR model must be modified to account for the difference in structure of lithium-like and hydrogenic ions considered previously. The same system of state labelling is used: the ground state ($2s$) is labelled 1, the first excited state ($2p$) is labelled 2, the second excited state ($3s$) is labelled 3, and so on in order of increasing level energy. The derivation of the equations of the CR plasma model for lithium-like ions then parallels that given previously for hydrogenic ions.

The time evolution of the population density of level p in an optically thin plasma is given by (13) as before. The steady state (SS) solution to the set of coupled first order differential equations (13) is obtained as before from (14). However, the quasi-steady state (QSS) solution must be modified to account for the small energy separation of the ground and the first excited states as compared to that of the first and the second excited states, as this is particularly significant for ions with large values of Z such as C IV, N V, and O VI. As a result of this, the population density of the first excited state (level 2) is very much larger than that of the other excited states, and it may even be comparable to that of the ground state.

Consequently, the QSS solution is modified by using a method similar to the one developed by Bates et al. [22] to describe hydrogenic plasmas optically thick toward the lines of the Lyman series. The normalized population density of level p is expressed as a function of the ground and the first excited state population densities:

$$\rho_p = r_p^{(0)} + r_p^{(1)}\, \rho_1 + r_p^{(2)}\, \rho_2 \tag{27}$$

where $3 \leq p \leq r$ and $r_p^{(0)}$, $r_p^{(1)}$ and $r_p^{(2)}$ are the population coefficients of level p. The QSS solution is obtained when the population densities of the second and higher excited states are in equilibrium with the population densities of the ground and the first excited states which, in general, are not in equilibrium. We then have $\dot{\rho}_1(t) \neq 0$, $\dot{\rho}_2(t) \neq 0$, and $\dot{\rho}_{p\geq 3}(t) = 0$.

Substituting the solution (27) in the system of equations (13), and using the last condition above, we obtain a set of equations of the form

$$a_p + b_p\, \rho_1 + c_p\, \rho_2 = 0\,;\; p = 3, 4, ..., r \,. \tag{28}$$

For arbitrary values of ρ_1 and ρ_2, the general solution of (28) is given by $a_p = 0$, $b_p = 0$, and $c_p = 0$. We must also impose the limiting conditions corresponding to the values of $p = 1, 2$ and $p > r$ on the population coefficients $r_p^{(0)}$, $r_p^{(1)}$ and $r_p^{(2)}$: $r_1^{(0)} = 0$, $r_1^{(1)} = 1$, $r_1^{(2)} = 0$; $r_2^{(0)} = 0$, $r_2^{(1)} = 0$, $r_2^{(2)} = 1$; $r_{p>r}^{(0)} = 1$, $r_{p>r}^{(1)} = 0$, $r_{p>r}^{(2)} = 0$. This last condition has already been applied to derive the system of equations (13).

Using these conditions, we obtain three sets of $r-2$ equations which are solved for the population coefficients $r_p^{(0)}$, $r_p^{(1)}$ and $r_p^{(2)}$ respectively:

$$\begin{aligned} &\sum_{q=3}^{p-1} \mathcal{A}_{pq}\, r_q^{(0)} - \mathcal{B}_p\, r_p^{(0)} + \sum_{q=p+1}^{r} C_{pq}\, r_q^{(0)} \\ &= -\frac{1}{Z_p}\left(\alpha_p\, n_e + \beta_p\right) - \sum_{q=r+1}^{s} \left(C_{p\to q}\, n_e + \frac{Z_q}{Z_p}\, A_{q\to p} \right) ; \end{aligned} \tag{29}$$

$$\sum_{q=3}^{p-1} \mathcal{A}_{pq}\, r_q^{(1)} - \mathcal{B}_p\, r_p^{(1)} + \sum_{q=p+1}^{r} C_{pq}\, r_q^{(1)} = -F_{p\to 1}\, n_e \,; \tag{30}$$

$$\sum_{q=3}^{p-1} \mathcal{A}_{pq}\, r_q^{(2)} - \mathcal{B}_p\, r_p^{(2)} + \sum_{q=p+1}^{r} C_{pq}\, r_q^{(2)} = -F_{p\to 2}\, n_e \tag{31}$$

where

$$\mathcal{A}_{pq} = F_{p\to q}\, n_e \,; \tag{32}$$

$$\mathcal{B}_p = \left(\sum_{q=1}^{p-1} F_{p\to q} + S_p + \sum_{q=p+1}^{s} C_{p\to q} \right) n_e + \sum_{q=1}^{p-1} A_{p\to q} \,; \tag{33}$$

$$C_{pq} = C_{p\to q}\, n_e + \frac{Z_q}{Z_p}\, A_{q\to p}\,;\; p = 3, 4, ..., r \,. \tag{34}$$

From the population coefficients $r_p^{(0)}$, $r_p^{(1)}$ and $r_p^{(2)}$, the population densities n_p can be calculated for any value of n_1 and n_2 from

$$\begin{aligned} &n_p = Z_p\, n_i\, n_e\, r_p^{(0)} + \frac{Z_p}{Z_1}\, n_1\, r_p^{(1)} + \frac{Z_p}{Z_2}\, n_2\, r_p^{(2)} \,; \\ &\quad p = 3, 4, ..., r \end{aligned} \tag{35}$$

where n_i is the ionic density. The time evolution of the population densities of the ground state and the first excited state, n_1 and n_2 respectively, can be obtained by substituting (27) and the population coefficients $r_p^{(0)}$, $r_p^{(1)}$ and $r_p^{(2)}$ into (13) with $p = 1$ and $p = 2$. We then get the two coupled first order differential equations

$$\begin{aligned} \dot{n}_1 &= -S_1^{CR}\, n_e\, n_1 + M_{21}^{CR}\, n_e\, n_2 + \alpha_1^{CR}\, n_e\, n_i \\ \dot{n}_2 &= -S_2^{CR}\, n_e\, n_2 + M_{12}^{CR}\, n_e\, n_1 + \alpha_2^{CR}\, n_e\, n_i \end{aligned} \tag{36}$$

where

$$\begin{aligned} S_1^{CR} &= S_1 + \sum_{q=2}^{s} C_{1\to q} - \\ &- \frac{1}{n_e\, Z_1} \sum_{q=3}^{s} \left(F_{q\to 1}\, n_e + A_{q\to 1} \right) Z_q\, r_q^{(1)} \,; \end{aligned} \tag{37}$$

$$S_2^{CR} = S_2 + F_{2\to1} + \frac{1}{n_e} A_{2\to1} + \sum_{q=3}^{s} C_{2\to q} - \\ -\frac{1}{n_e Z_2} \sum_{q=3}^{s} \left(F_{q\to2}\, n_e + A_{q\to2}\right) Z_q\, r_q^{(2)}\,; \tag{38}$$

$$\alpha_1^{CR} = \alpha_1\, n_e + \beta_1 + \sum_{q=3}^{s} \left(F_{q\to1}\, n_e + A_{q\to1}\right) Z_q\, r_q^{(0)}\,; \tag{39}$$

$$\alpha_2^{CR} = \alpha_2\, n_e + \beta_2 + \sum_{q=3}^{s} \left(F_{q\to2}\, n_e + A_{q\to2}\right) Z_q\, r_q^{(0)}\,; \tag{40}$$

$$M_{21}^{CR} = F_{2\to1} + \frac{1}{n_e} A_{2\to1} \\ +\frac{1}{n_e Z_2} \sum_{q=3}^{s} \left(F_{q\to1}\, n_e + A_{q\to1}\right) Z_q\, r_q^{(2)}\,; \tag{41}$$

$$M_{12}^{CR} = C_{1\to2} + \frac{1}{n_e Z_1} \sum_{q=3}^{s} \left(F_{q\to2}\, n_e + A_{q\to2}\right) Z_q\, r_q^{(1)}\,. \tag{42}$$

The coefficients S_1^{CR}, S_2^{CR} and α_1^{CR}, α_2^{CR} are similar to the hydrogenic collisional-radiative ionization rate coefficient S_{CR} (23) and recombination rate coefficient α_{CR} (24) respectively. The coefficients M_{21}^{CR} and M_{12}^{CR} have no hydrogenic counterparts. The collisional-radiative rate coefficient M_{21}^{CR} expresses the recombination which occurs in the ground state due to the neighbouring first excited state and vice versa for the collisional-radiative rate coefficient M_{12}^{CR}.

The general solution of the coupled system of equations (36) can be written as

$$n_j(t) = n_j^{SS} + n_j^{(+)} e^{-\lambda^{(+)} t} - n_j^{(-)} e^{-\lambda^{(-)} t} \tag{43}$$

where $j = 1$ or 2,

$$\lambda^{(\pm)} = \frac{n_e}{2} \left(S_1^{CR} + S_2^{CR} \pm \\ \pm \sqrt{\left(S_1^{CR} - S_2^{CR}\right)^2 + 4\, M_{12}^{CR}\, M_{21}^{CR}} \right), \tag{44}$$

$$n_j^{SS} = \frac{K_j^{SS}}{\lambda^{(+)}\, \lambda^{(-)}}\,, \tag{45}$$

$$n_j^{(\pm)} = \frac{n_j(t=0)\, \lambda^{(\pm)2} - K_j \lambda^{(\pm)} + K_j^{SS}}{\lambda^{(\pm)}\, \left(\lambda^{(+)} - \lambda^{(-)}\right)}\,, \tag{46}$$

$$K_1^{SS} = n_e^2\, n_i \left(\alpha_1^{CR}\, S_2^{CR} + \alpha_2^{CR}\, M_{21}^{CR}\right), \tag{47}$$

$$K_2^{SS} = n_e^2\, n_i \left(\alpha_2^{CR}\, S_1^{CR} + \alpha_1^{CR}\, M_{12}^{CR}\right), \tag{48}$$

$$K_1 = n_e \left(\alpha_1^{CR}\, n_i + S_2^{CR}\, n_1(t=0) + M_{21}^{CR}\, n_2(t=0)\right), \tag{49}$$

$$K_2 = n_e \left(\alpha_2^{CR}\, n_i + S_1^{CR}\, n_2(t=0) + M_{12}^{CR}\, n_1(t=0)\right). \tag{50}$$

The steady state population densities, which are obtained in the limit as $t \to \infty$, are explicitly given by

$$n_1^{SS} = \frac{\alpha_1^{CR}\, S_2^{CR} + \alpha_2^{CR}\, M_{21}^{CR}}{S_1^{CR}\, S_2^{CR} - M_{12}^{CR}\, M_{21}^{CR}}\, n_i\,; \tag{51}$$

$$n_2^{SS} = \frac{S_1^{CR}\, \alpha_2^{CR} + \alpha_1^{CR}\, M_{12}^{CR}}{S_1^{CR}\, S_2^{CR} - M_{12}^{CR}\, M_{21}^{CR}}\, n_i\,. \tag{52}$$

4.6 Calculation of collisional and radiative rate coefficients

The results of the modelling calculations depend to a large extent on the accuracy of the collisional and radiative rate coefficients used in the CR model. The collisional rate coefficients R_n are obtained by integrating the cross-sections σ_n of the collisional processes over the free electron velocity distribution, $f(v)$:

$$R_n(T) = \int_v \sigma_n(v)\, v\, f(v)\, \mathrm{d}v\,. \tag{53}$$

For a Maxwellian velocity distribution of the free electrons, we have

$$f(v)\, \mathrm{d}v = \frac{4}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{3/2} v^2\, \exp(-mv^2/2kT)\, \mathrm{d}v\,. \tag{54}$$

The cross-section values are obtained from experimental data, where available, and from various model and theoretical calculations that are usually fitted to semi-empirical expressions. We briefly review the expressions that have been found to be useful in CR model calculations [40].

The spontaneous transition probabilities from an upper state n to a lower state n' are given, within the electric dipole approximation, by the Einstein probability coefficient [31]

$$A_{n\to n'} = \frac{8\pi^2 e^2}{mc^3}\, \nu_{nn'}^2\, \frac{\omega_{n'}}{\omega_n}\, f_{n'\to n} \tag{55}$$

where ω_n and $\omega_{n'}$ are the statistical weights of levels n and n' respectively, $\nu_{nn'}$ is the frequency of the photon emitted as a result of the transition and $f_{n'\to n}$ is the absorption oscillator strength for the $n' \to n$ transition. The oscillator strengths can be evaluated exactly for hydrogenic ions using hypergeometric functions. Average lifetime of hydrogenic levels can be calculated from the asymptotic expression given by Millette [32]. For other elements, oscillator strengths for allowed and forbidden transitions can be evaluated using various approximate theoretical methods.

The cross-section for collisional excitation of the optically allowed transition $n' \to n$ by electron impact is given by [40]

$$\sigma_{n'\to n}(u) = 4\pi\, a_0^2\, \frac{f_{n'\to n}}{E_{n'n}^2}\, \alpha_{n'n}\, \frac{u - \phi_{n'n}}{u^2}\, \ln(1.25\, \beta_{n'n}\, u) \tag{56}$$

where $E_{n'n}$ is the threshold energy for the excitation of the $n' \rightarrow n$ transition in Rydbergs, $u = E/E_{n'n}$ is the energy of the impacting electron E in threshold units, $f_{n' \rightarrow n}$ is the absorption oscillator strength for the $n' \rightarrow n$ transition, a_0 is the Bohr radius and $\alpha_{n'n}$, $\beta_{n'n}$ and $\phi_{n'n} \leq 1$ (equal to 1 for atoms) are fit parameters depending on the transition.

The cross-section for collisional excitation of the optically forbidden transition $n' \rightarrow n$ by electron impact is given by [40]

$$\sigma_{n' \rightarrow n}(u) = 4\pi\, a_0^2 \left(\frac{n'}{n}\right)^3 \frac{\alpha_{n'n}}{E_{n'n}^2} \frac{u - \phi_{n'n}}{u^2} \tag{57}$$

where $E_{n'n}$ is the threshold energy for the excitation of the $n' \rightarrow n$ transition in Rydbergs, $u = E/E_{n'n}$ is the energy of the impacting electron E in threshold units, a_0 is the Bohr radius and $\alpha_{n'n}$ and $\phi_{n'n}$ are fit parameters depending on the transition.

The collisional de-excitation rate coefficients are obtained from the collisional excitation rate coefficients by the principle of detailed balancing as given by (12).

The collisional ionization cross-section from state n by electron impact is given by [33, 40]

$$\sigma_n(u) = 2.66\, \pi a_0^2 \left(\frac{I_1^H}{I_n}\right)^2 \xi_n \frac{u - 1}{u^2} \ln(1.25\beta_n\, u) \tag{58}$$

where $I_1^H = E_1^H$ is the ionization energy of the hydrogen atom in its ground state, $I_n = E_n$ is the ionization energy of the atom or ion in state n, $u = E/I_n$ is the kinetic energy of the incident electron in units of the threshold energy for ionization from state n, ξ_n is the number of equivalent electrons in state n and β_n is a correction (fit) factor of order unity. To obtain the correct threshold law, β_n must be larger than 0.8.

The three-body recombination rate coefficients are obtained from the collisional ionization rate coefficients by the principle of detailed balancing.

The radiative recombination rate coefficients can be obtained from the photo-ionization rate coefficients by the principle of detailed balancing. The available experimental and calculated photo-ionization data are fitted to a semi-empirical function of the form [40]

$$a(u) = \frac{C}{u^p} \left[1 + \frac{b_1}{u} + \frac{b_2}{u^2} + \cdots + \frac{b_m}{u^m}\right] \tag{59}$$

where u is the energy of the incident photon in threshold energy units, and C and $b_k,\ k = 1, ..., m$ are fit parameters. The parameters p and m are restricted to the range of values $0 \leq p \leq 5$ and $1 \leq m \leq 9$, and p is assigned half-integral values to simplify and facilitate the evaluation of the rate coefficient integrals.

5 Laser action in Wolf-Rayet stars

The strength of an inversely populated transition $q \rightarrow p$ ($p < q$) can be characterized by the fractional gain per unit distance, α. At the centre of a Doppler-broadened line, it is given by the following expression [34, p, 23]:

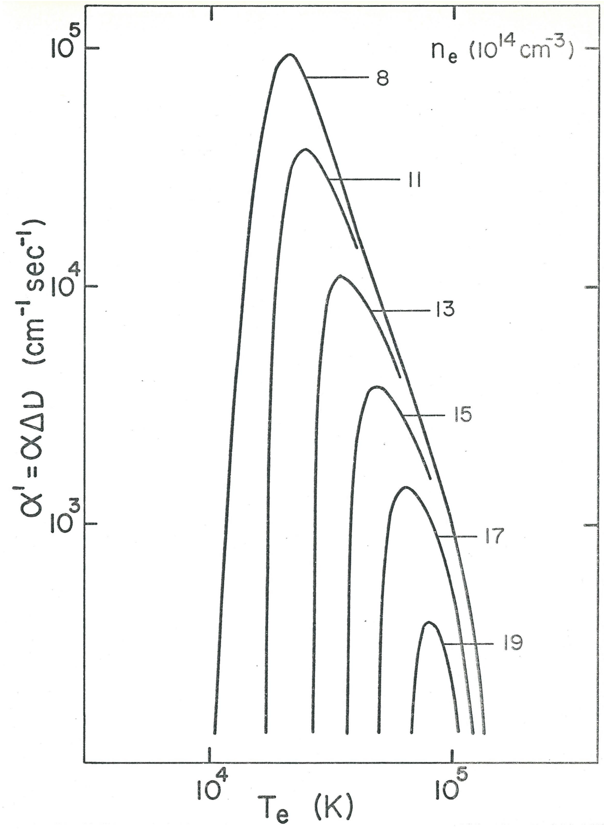


Fig. 2: Typical α' versus T_e plot for the $6f \rightarrow 5d$ transition of C IV [40, p. 249].

$$\alpha = \sqrt{\frac{\ln 2}{\pi}} \left(\frac{\omega_q A_{q \rightarrow p}}{4\pi}\right) \frac{P \lambda_0^2}{\Delta\nu} \tag{60}$$

where λ_0 is the centre wavelength of the transition, $\Delta\nu$ is the linewidth, ω_q is the statistical weight of level q, $A_{q \rightarrow p}$ is the Einstein probability coefficient for spontaneous transition from level q to p, and [35]

$$P = \frac{n_q}{\omega_q} - \frac{n_p}{\omega_p}\,. \tag{61}$$

P is a measure of the population inversion and, for laser action to be operative, $P > 0$. α is related to the intensity of a plane wave at λ_0 by the equation

$$I = I_0\, e^{\alpha L} \tag{62}$$

where L is the length over which gain occurs. To be able to compare various transitions without needing to specify the linewidth $\Delta\nu$, we define a quantity α' given by

$$\alpha' = \alpha\, \Delta\nu \tag{63}$$

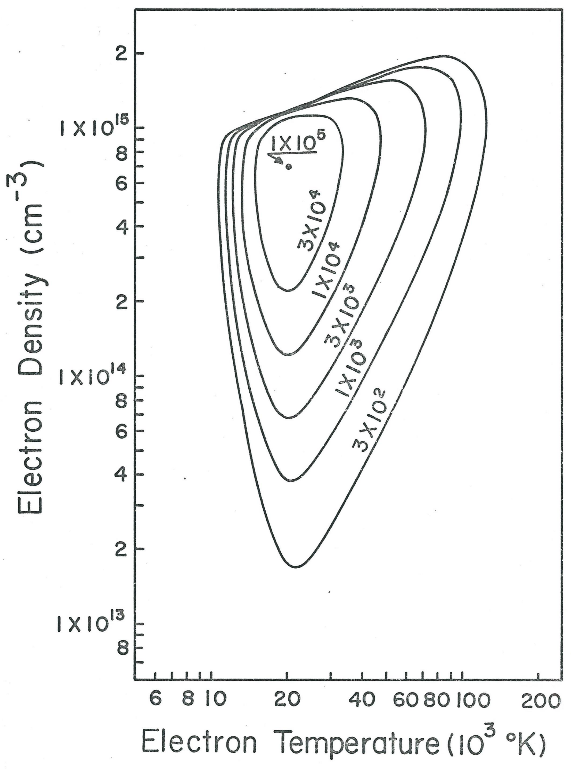


Figure 9.13 - n_e-T_e diagram for the 6f->5d transition of C IV (λ4646).

Fig. 3: Typical $n_e - T_e$ diagram showing laser gain equi-α' contours in cm^{-1} s^{-1} for the $6f \to 5d$ transition of C IV [40, p. 257].

where α is given by (60).

Model calculations starting from an initial element number density of 10^{14} cm^{-3} are performed for a grid of n_e and T_e values. The inversion is displayed on n_e, T_e plots ($n_e - T_e$ diagrams) showing contours of equal P or α' (equi-α' contours). Fig. 2 shows a typical variation of α' versus T_e for inversely populated transition $6f \to 5d$ of C IV. Fig. 3 shows a typical $n_e - T_e$ diagram with equi-α' contours for inversely populated transition $6f \to 5d$ of C IV.

On a three-dimensional plot with α' as the third axis perpendicular to both the n_e and T_e axes, the diagram would appear as a triangular pyramidal-shaped mountain with a very steep slope on the high-n_e side, a steep slope on the low-T_e side, and a gradual slope on the low-n_e, high-T_e side. Strong population inversion thus occurs only within a narrow range of values of n_e and T_e, and each transition has its own region of strong population inversion. This provides a means to classify Wolf-Rayet star parameters from their spectra.

Calculations of population inversions in astronomical plasmas cooled by adiabatic expansion have been performed on ions observed in WR stars by the following investigators. Varshni and Lam [37–39] investigated population inversions in the hydrogen-like He II ion for line λ4686 resulting from the transition 4 → 3 in He II.

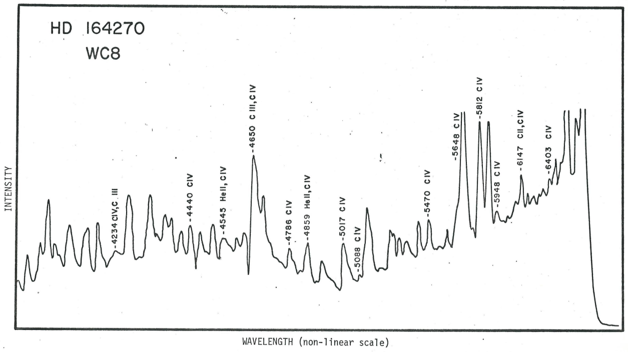


Fig. 4: Spectrum of the WC8 star HD 164270 from [36].

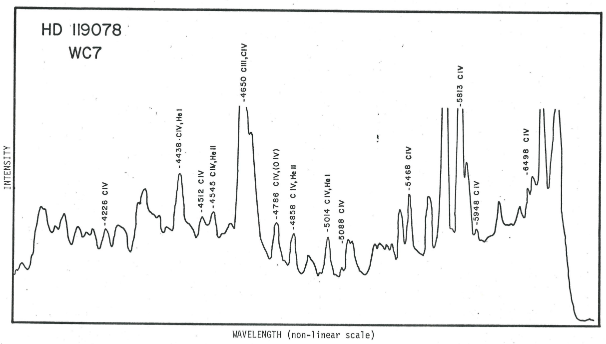


Fig. 5: Spectrum of the WC7 star HD 119078 from [36].

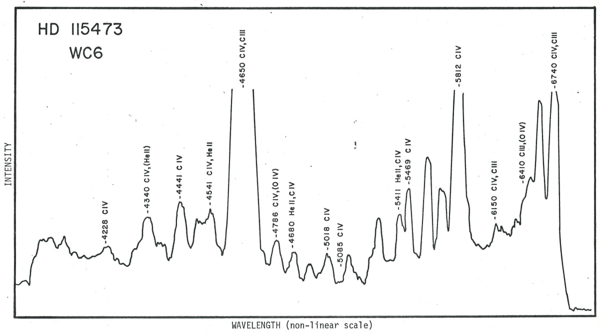


Fig. 6: Spectrum of the WC6 star HD 115473 from [36].

Millette [40] analyzed population inversions in the lithium-like ions C IV, N V and O VI. Population inversions were found to occur in many of the transitions. C IV transitions giving rise to emission lines in the visible region of the spectrum, specifically line λ4650 resulting from transitions between levels 6 → 5 in C IV, were investigated. The C IV $\lambda\lambda$4646, 4658 lines arising from the $6f \to 5d$ and $6g \to 5f$ transitions respectively, were found to be strongly inverted allowing laser action in plasmas cooled by adiabatic expansion.

The model calculations provide an understanding of the unusual strength of the C IV λ4650 emission line in the WC category of Wolf-Rayet stars, as seen in Fig. 4, Fig. 5 and Fig. 6, which shows the $\lambda 4650$ line becoming more and more prominent in going from a category WC8 to a WC6 Wolf-Rayet star. The lines in WC8 WR stars are relatively sharp, becoming wider and brighter in WC7 WR stars, and even wider and brighter in WC6 WR stars, indicating increasing speed of ejection and increasing laser action.

Varshni and Nasser [41,42] investigated population inversions in He I and in helium-like C III. Four transitions were investigated in the visible region of He I, $\lambda 7281\ 3^1S \rightarrow 2^1P$, $\lambda 6678\ 3^1D \rightarrow 2^1P$, $\lambda 5047\ 4^1S \rightarrow 2^1P$ and $\lambda 4922\ 4^1D \rightarrow 2^1P$, of which observationl evidence is available for $\lambda 7281$ and $\lambda 6678$ in WR stars. Two transitions showed appreciable population inversion in the visible region of C III: $\lambda 4650\ 2s3p\,^3S \rightarrow 2p3p\,^3S$ and $\lambda 5263\ 2p3p\,^3S \rightarrow 2p3s\,^3P^0$.

Millette [40] provides a detailed roadmap to calculate population inversions in hydrogenic and in the lithium-like ions N V and O VI, in addition to C IV.

6 Laser action in Quasi-Stellar Objects

The physical process of population inversions in expanding stellar atmospheres led Varshni to formulate his Plasma Laser Star (PLS) model as an explanation of the spectra of Wolf-Rayet stars and Quasi-Stellar Objects [43–48]. Radio astronomy first detected QSOs in the 1950s as anomalous objects with unexplained properties. QSO 3C 273 was the first radio source quasar for which an optical counterpart was identified in 1963. Its spectrum consisted of one strong emission line and one medium to weak strength line ($\lambda 5637$, $\lambda 7588$).

QSOs were named quasi-stellar because they look like stars, if not typical stars. In particular QSO spectra are dominated by a small number of very intense and wide lines that could not be readily identified with common elements. In particular, there was a lack of the expected hydrogen Lyman lines, a typical marker in most spectra. This likely provided the impetus for Schmidt [49] to assume that the observed lines in 3C 273 were the Hα and Hβ lines, red-shifted to their observed wavelength in the spectrum. This quickly became the standard approach, and ever since, astronomy and cosmology have been transformed, with everything looking like red-shifted objects, even if those red-shifts are superluminal.

Luckily, this possibility did not exist when Wolf-Rayet stars were first discovered in 1867 by astronomers Charles Wolf and George Rayet at the Paris Observatory, otherwise we would be facing an even more confusing puzzle, as hydrogen emission lines are not present in WR spectra either. As chance would have it, WR stars were investigated as stellar objects, which allowed us to eventually determine the presence of laser action in WR stellar atmospheres, which is the same process that is operating in QSO stellar atmospheres.

Banerji and Bhar [50–52] have compared the (unshifted) spectral lines of 633 QSOs discovered till August 1976, assuming they are generated by a population inversion process similar to that operating in WR stars instead of red shifts, with the laser transitions observed in laboratories till April 1976 [53]. They found that 88% of the QSO lines agreed to within 10 Å with the laser lines and 94% agreed to within 20 Å. Their assumption that QSOs are early-type stars with temperatures in the range 104–105 K implied spectral lines with asymmetric shapes and large broadening leading to errors in measurement of up to 20 Å. They pointed out the similarities between the spectra of QSOs and those of Wolf-Rayet stars, with both deficient in hydrogen. They proposed that the absorption lines of QSOs are produced in the expanding stellar atmosphere, so that they are violet-shifted as in WR stars. Under this model, they showed that 54 of 55 narrow absorption lines in QSO Q 1246-057 can be explained by assuming an average velocity of absorbing ions of 500 km/s.

Taking Quasi-Stellar Objects to be local stellar objects instead of distant galactic objects eliminates the problems associated with their currently accepted cosmological interpretation: energy source, superluminal velocities, optical variability, quasar proper motions [54, 55], quasar binary systems [56,57], naked (no nebulosity) quasars, *etc*. The properties of QSOs are similar to those of WR stars and, as stars, those are easily explainable in terms of commonly known physical processes.

7 A new star type Q and the Hertzsprung-Russell diagram

We consider the implications of Quasi-Stellar Objects as stellar objects. We need to first be more specific about the terminology used: we use the term *quasar* to refer specifically to the cosmological interpretation of Quasi-Stellar Objects, while we use the term *QSO* to refer to the stellar interpretation of Quasi-Stellar Objects. We introduce a new star type to denote QSOs: stars of type Q, similar to the Wolf-Rayet stars which are denoted as stars of type W.

The Hertzsprung-Russell diagram is extended beyond the stars of type O B towards more massive and hotter stars of type Q and W. The main sequence starts with Q W O B, followed by the standard A F G K M types of the rest of the sequence. As one moves towards star type Q, the stars become increasingly more massive, of higher temperature, with higher speeds of stellar atmosphere ejection and population inversions, with their emission spectra increasingly dominated by the lasing emission lines.

Significant work has been performed on the analysis of WR stars to understand their classification and evolution. WR stars are known to be hot, luminous objects, representative of the late stage of evolution of massive O stars. The details have been worked out over the last forty years [2, 10, 58–65] with the analysis of Wolf-Rayet stars in the Magellanic Clouds dwarf satellite galaxies of the Milky Way providing valuable information. A similar effort is required to understand the classification and evolution of stars of type Q, with the identification of unrecognized representatives in our galaxy and in the Magellanic Clouds an important step [55, 66].

8 Discussion and conclusion

In this paper, we have reconsidered the little-known but critically important physical process of laser action occurring in the stellar atmospheres of Wolf-Rayet stars and, by extension, of QSOs. We have reviewed the model used for hydrogenic and lithium-like ions in the Collisional-Radiative (non-LTE) model used to calculate the ionic energy level populations, and the existing results for He I, He II, C III and C IV. We have noted the availability of a detailed roadmap in [40] to carry out similar calculations for the lithium-like ions of interest N V and O VI.

We have reviewed the details of laser action in Wolf-Rayet stars. We have considered the historical bifurcation that resulted in the red-shift model of quasar spectra and its cosmological roots. We have also considered the evidence for the presence of laser action in QSOs as in Wolf-Rayet stars, and how taking QSOs to be local stellar objects instead of distant galactic objects eliminates the problems associated with the currently accepted cosmological interpretation.

We have introduced the terminology *quasar* to refer specifically to the cosmological interpretation of Quasi-Stellar Objects and *QSO* to refer to the stellar interpretation of Quasi-Stellar Objects. We have introduced a new star type Q for QSOs, similar to the star type W for Wolf-Rayet stars. We have expanded the Hertzsprung-Russell diagram to include more massive and hotter stars of type Q and W beyond the stars of type O B. The main sequence thus starts with stars of type Q W O B, followed by the standard types A F G K M of the rest of the sequence. Finally, we have noted the effort that will be required to understand the classification and evolution of stars of type Q, as has been achieved for Wolf-Rayet stars.

Received on December 27, 2020

References

1. Menzel D. H. Laser Action in non-LTE Atmospheres. International Astronomical Union Colloquium, Volume 2: Spectrum Formation in Stars With Steady-State Extended Atmospheres, 1970, 134–137.
2. Crowther P. A. Physical Properties of Wolf-Rayet Stars. *Annu. Rev. Astron. Astrophys.*, 2007, v. 45, 177–219. arXiv: astro-ph/0610356v2.
3. Mihalas D. Stellar Atmospheres, 2^{nd} ed. W. H. Freeman and Co., San Francisco, 1978.
4. van der Hucht K. A. The VIIth catalogue of galactic Wolf-Rayet stars. *New Astronomy Reviews*, 2001, v. 45, 135–232.
5. van der Hucht K. A. New Galactic Wolf-Rayet stars, and candidates (Research Note), An Annex to The VIIth Catalogue of Galactic Wolf-Rayet Stars. *Astronomy and Astrophysics*, 2006, v. 458, 453–459.
6. Galactic Wolf Rayet Catalogue. V1.25, www.pacrowther.staff.shef.ac.uk/WRcat/, Aug. 2020.
7. Beals C. S. On the Nature of Wolf-Rayet Emission. *Mon. Not. Royal Ast. Soc.*, 1929, v. 90, 202–212.
8. Beals C. S. The Contours of Emission Bands in Novae and Wolf-Rayet Stars. *Mon. Not. Royal Ast. Soc.*, 1931, v. 91, 966.
9. Mihalas D. and Weibel-Mihalas B. Foundations of Radiation Hydrodynamics, corr. ed. Dover Publications, New York, 1999, pp. 627–645.
10. Abbott D. C. and Conti P. S. Wolf-Rayet Stars. *Ann. Rev. Astron. Astrophys.*, 1987, v. 25, 113–150.
11. Castor J. I. Spectral Line Formation in Wolf-Rayet Envelopes. *Mon. Not. R. Astr. Soc.*, 1970, v. 149, 111–127.
12. Castor J. I. and Van Blerkom D. Excitation of He II in Wolf-Rayet Envelopes. *Astrophys. J.*, 1970, v. 161, 485–502.
13. Castor J. I. and Nussbaumer H. On the Excitation of C III in Wolf-Rayet Envelopes. *Mon. Not. R. Astr. Soc.*, 1972, v. 155, 293–304.
14. Gudzenko L. L., Shelepin L. A. Negative Absorption in a Nonequilibrium Hydrogen Plasma. *Zh. Eksp. Teor. Fiz.*, 1963, v. 45, 1445–1449. *Sov. Phys. JETP*, 1964, v. 18, 998–1000.
15. Gudzenko L. L., Shelepin L. A. Amplification in Recombination Plasma. *Dokl. Akad. Nauk. SSSR*, 1965, v. 160, 1296–1299. *Sov. Phys. - Dokl.*, 1965, v. 10, 147.
16. Gudzenko L. L., Filippov S. S., Shelepin L. A. Rapid Recombination of Plasma Jets. *Zh. Eksp. Teor. Fiz.*, 1966, v. 51, 1115–1119. *Sov. Phys. JETP*, 1967, v. 24, 745–748.
17. Gudzenko L. I., Mamachun A. T., and Shelepin L. A. *Zh. Tekh. Fiz.*, 1967, v. 37, 833. *Sov. Phys. - Tech. Phys.*, 1967, v. 12, 598.
18. Gudzenko L. L., Shelepin L. A., Yakovlenko S. I. Amplification in recombining plasmas (plasma lasers). *Usp. Fiz. Nauk*, 1974, v. 114, 457. *Sov. Phys. - Usp.*, 1975, v. 17, 848.
19. Gol'dfarb V. M., Il'ina E. V., Kostygova I. E., Luk'yanov G. A. and Silant'ev, V. A. *Opt. Spektrosk.*, 1966, v. 20, 1085. *Optics and Spectrosc.*, 1969, v. 20, 602.
20. Gol'dfarb V. M., Il'ina E. V., Kostygova I. E. and Luk'yanov G. A. *Opt. Spektrosk.*, 1969, v. 27, 204. *Optics and Spectrosc.*, 1969, v. 27, 108.
21. Bates D. R., Kingston A. E. and McWhirter R. W. P. *Proc. Roy. Soc.*, 1962, v. A267, 297.
22. Bates D. R., Kingston A. E. and McWhirter R. W. P. *Proc. Roy. Soc.*, 1962, v. A270, 155.
23. Bates D. R. and Kingston A. E. *Planet. Space Sci.*, 1963, v. 11, 1.
24. McWhirter R. W. P. and Hearn A. G. *Proc. Phys. Soc.*, 1963, v. 82, 641.
25. Drawin H. W. and Emard F. Instantaneous Population Densities of the Excited Levels of Hydrogen Atoms, Hydrogen-like Ions, and Helium Atoms in Optically Thin and Thick Non-LTE Plasmas. Euratom-C.E.A. Report No. EUR-CEA-FC-534, 1970.
26. Gordiets B. F., Gudzenko L. I. and Shelepin L. A. Relaxation processes and amplification of radiation in a dense plasma. *Zh. Eksp. Teor. Fiz.*, 1968, v. 55, 942. *Sov. Phys. - JETP*, 1969, v. 28, 489.
27. Norcross D. W. and Stone P. M. *J. Quant. Spectrosc. Radiat. Transfer*, 1968, v. 8, 655.
28. Unsöld A. The New Cosmos, 2^{nd} edition. Springer-Verlag, Berlin, 1977.
29. Griem H. R. Plasma Spectroscopy. McGraw-Hill, New York, 1964.
30. House L. L. *Astrophys. J. Suppl.*, 1964, v. 8, 307.
31. Menzel D. H. and Pekeris C. L. *Mon. Not. Roy. Astron. Soc.*, 1935, v. 96, 77.
32. Millette P. A. and Varshni Y. P. New Asymptotic Expression for the Average Lifetime of Hydrogenic Levels. *Can. J. Phys.*, 1979, v. 57 (3), 334–335.
33. Drawin H. W. Collision and Transport Cross-Sections. Euratom-C.E.A. Report No. EUR-CEA-FC-383, 1966.
34. Willett C. S. Gas Lasers: Population Inversion Mechanisms. Pergamon Press, New York, 1974.

35. Lengyel B. A. Introduction to Laser Physics. John Wiley, New York, 1966.
36. Smith H. J. Ph.D. Thesis, Dept. of Astronomy, Harvard University, Cambridge, 1955.
37. Varshni Y. P., Lam C. S. Emission Line 4686 in the Quasi-Stellar Objects. *J. Roy.Astron. Soc. Canada*, 1974, v. 68, 264.
38. Varshni Y. P., Lam C. S. Laser Action in Stellar Envelopes. *Bull. Amer. Astron. Soc.*, 1975, v. 7, 551.
39. Varshni Y. P., Lam C. S. Laser Action in Stellar Envelopes. *Astrophys. Space Sci.*, 1976, v. 45, 87.
40. Millette P. A. Laser Action in C IV, N V, and O VI Plasmas Cooled by Adiabatic Expansion. University of Ottawa, Ottawa, ON, 1980. LaTeX typeset, researchgate.net/publication/283014713, 2015.
41. Nasser R. M. Population Inversion in He I and C III in Recombining Plasmas. University of Ottawa, Ottawa, ON, 1986.
42. Varshni, Y. P., Nasser R. M. Laser Action in Stellar Envelopes II. He I. *Astrophys. Space Sci.*, 1986, v. 125, 341.
43. Varshni Y. P. Laser Action in Quasi-Stellar Objects? *Bull. Amer. Phys. Soc.*, 1973, v. 18, 1384.
44. Varshni Y. P. No Redshift in Quasi-Stellar Objects. *Bull. Amer. Astron. Soc.*, 1974, v. 6, 213.
45. Varshni Y. P. The Redshift Hypothesis and the Plasma Laser Star Model for Quasi-Stellar Objects. *Bull. Amer. Astron. Soc.*, 1974, v. 6, 308.
46. Varshni Y. P. Alternative Explanation for the Spectral Lines Observed in Quasars. *Astrophys. Space Sci.*, 1975, v. 37, L1.
47. Varshni Y. P. Electron Density in the Emission-Line Region of Wolf-Rayet Stars. *Astrophys. Space Sci.*, 1978, v. 56, 385.
48. Varshni Y. P. The Physics of Quasars. *Phys. Canada*, 1979, v. 35, 11.
49. Schmidt M. 3C 273: A Star-Like Object with Large Red-Shift. *Nature*, 1963, v. 197, 1040.
50. Banerji S. and Bhar G. C. Plasma laser star model of QSOs. *Astrophysics and Space Science*, 1978, v. 56, 443–451.
51. Banerji S. and Bhar G. C. Analysis of the plasma laser star model of QSOs. *Astrophysics and Space Science*, 1979, v. 61, 337–347.
52. Banerji S., Bhar G. C. and Mukherji P, K. Are QSOs local objects? I. A new interpretation of the emission and absorption spectra of a few QSOs. *Astrophysics and Space Science*, 1982, v. 87, 217–236.
53. Willett, C. S. Laser Lines in Atomic Species. In: Progress in Quantum Electronics, Vol. 1, Part 5, Pergamon, 1971.
54. Luyten W. J. A Search for Faint Blue Stars. Paper 50, University of Minnesota Observatory, Minneapolis, 1969.
55. Varshni Y. P. Proper Motions and Distances of Quasars. *Speculations in Science and Technology*, 1982, v. 5 (5), 521–532.
56. Mortlock D. J., Webster R. L. and Francis P. J. Binary Quasars. *Mon. Not. R. Astron. Soc.*, 1999, v. 309, 836–846.
57. Hennawi J. F., et al. Binary Quasars in the Sloan Digital Sky Survey: Evidence for Excess Clustering on Small Scales. *Astronomical Journal*, 2006, v. 131, 1–23.
58. Conti P. S. and Garmany C. D., De Loore C. and Vanbeveren D. The Evolution of Massive Stars: The Numbers and Distribution of O Stars and Wolf-Rayet Stars. *Astrophys. J.*, 1983, v. 274, 302–312.
59. Humphreys R. M. and Nichols M., Massey P. On the Initial Masses and Evolutionary Origins of Wolf-Rayet Stars. *Astron. J.*, 1985, v. 90, 101–108.
60. Conti P. S. and Vacca W. D. The Distribution of Massive Stars in the Galaxy: I. Wolf-Rayet Stars. *Astron. J.*, 1990, v. 100, 431–444.
61. Langer N., Hamann W.-R., Lennon M., Najarro F., Pauldrach A.Ẇ. A. and Puls J. Towards an understanding of very massive stars. A new evolutionary scenario relating O stars, LBVs [Luminous Blue Variables] and Wolf-Rayet stars. *Astron. Astrophys.*, 1994, v. 290, 819–833.
62. Crowther P. A., Smith L. J., Hillier D. J. and Schmutz W. Fundamental parameters of Wolf-Rayet stars. III. The evolutionary status of WNL [WN7 to WN9] stars. *Astron. Astrophys.*, 1995, v. 293, 427–445.
63. Hainich R., Rühling U., Todt H., Oskinova L. M., Liermann A., Gräfener G., Foellmi C., Schnurr O. and Hamann W.-R. The Wolf-Rayet stars in the Large Magellanic Cloud. A comprehensive analysis of the WN class. *Astronomy & Astrophysics*, 2014, v. 565, A27, 1-62.
64. Koesterke L. and Hamann W.-R. Spectral analyses of 25 Galactic Wolf-Rayet stars of the carbon sequence. *Astron. Astrophys.*, 1995, v. 299, 503–519.
65. Barlow M. J. and Hummer D. J. The WO Wolf-Rayet stars. *Symposium – International Astronomical Union*, 1982, v. 99, 387–392.
66. Varshni Y. P. O VI and He II Emission Lines in the Spectra of Quasars. *Astrophys. Space Sci.*, 1977, v. 46, 443.

Physics of Transcendental Numbers Meets Gravitation

Hartmut Müller

E-mail: hm@interscalar.com

Transcendental ratios of physical quantities can provide stability in complex dynamic systems because they inhibit the occurrence of destabilizing resonance. This approach leads to a fractal scalar field that affects any type of physical interaction and allows reformulating and resolving some unsolved tasks in celestial mechanics and astrophysics. We verify the model claims on the gravitational constants and the periods of orbital and rotational motion of the planets, planetoids and large moons of the solar system as well as the orbital periods of exoplanets and the gravitational constants of their stars.

Introduction

Despite the abundance of theoretical approaches engaged to explain the origin of gravitational interaction dealing with superstrings, chameleons or entropic forces [1], the community of physicists still expects compatibility for centuries: any modern theory must allow deriving Newton's law of universal gravitation as classic approximation. In the normal case of weak gravity and low velocities, also Einstein's field equations obey the correspondence principle.

Besides of nostalgia, what could be the reason of this condition? Newton's law of gravitation cannot be verified in the scale of the solar system, because the mass of a planet cannot be measured, and Kepler's laws of planetary motion do not compellingly require Newton's law of gravitation for their derivation. Moreover, Newton's theory of gravitation leads to inconsistencies already in the case of three interacting bodies.

It is a common belief that John Couch Adams and Urbain Le Verrier applying Newton's law of gravitation could predict the orbit and correct position of Neptune based on motions of Uranus. However, this is not exactly what they did.

Adopting the Titius-Bode law [2], Adams assumed the semi-major axis of Neptune being 37.25 AU; Le Verrier estimated 36.15 AU. The deviation from the correct value 30.07 AU is more than 20%. Adopting Pontécoulant's Théorie Analytique to his perturbation approach, Adams calculated an eccentricity of 0.1206; Le Verrier got 0.1076. The right value is 0.0086, a deviation of more than 1100%. Adams calculated the longitude of the perihelion being at 299°; Le Verrier arrived at 284° while the correct is 44°. Finally, applying Newton's law of gravitation, Adams estimated Neptune's mass with 1/6666 solar mass; Le Verrier calculated 1/9300. Actually, the ratio is 1/19300. Again, a deviation of > 200%. It is a miracle how with all these errors Le Verrier could guess the right longitude 326° of the current position of Neptune. Obviously, he was very lucky [3].

Kepler's laws of planetary motion cannot explain why the solar system has established the orbital periods 90560 days (Pluto), 60182 (Neptune), 30689 (Uranus), 10759 (Saturn), 4333 (Jupiter), 1682 (Ceres), 687 (Mars), 365 (Earth), 225 (Venus) and 88 days (Mercury), because there are infinitely many pairs of orbital periods and distances that fulfill Kepler's laws. Einstein's field equations do not reduce the theoretical variety of possible orbits, but increases it even more.

But now, after the discovery of thousands of exoplanetary systems, we can recognize that the current distribution of the planetary and lunar orbits in our solar system is not accidental. Many planets in the extrasolar systems like Trappist 1 or Kepler 20 have nearly the same orbital periods as the large moons of Jupiter, Saturn, Uranus and Neptune [4]. That's amazing, because Trappist 1 is 40 light years away from our solar system and Kepler 20 nearly 1000 light years [5, 6].

The question is, why they prefer similar orbital periods if there are infinite possibilities? Obviously, there are orbital periods preferred anywhere in the galaxy. Why these orbital periods are preferred? What makes them attractive?

Despite perturbation models and parametric optimization, the reality of planetary systems is still a theoretical problem. The notoriously high failure rate of interplanetary missions, flyby anomalies [7] and unexpected accelerations of spacecraft indicate a profound lack of understanding gravity.

In spiral galaxies, the orbiting of stars around their centers seems to strongly disobey both Newton's law of universal gravitation and general relativity. Recently, an 85% dark matter universe is required for saving the conventional paradigm.

Perhaps the concept of gravitation itself requires a revision. Obviously, it is not about details, but an important part of the hole is missing. For finding the missing part, let us go back to the roots of the idea of gravitation . . .

The empirical universality of free fall led ancient philosophers to the idea that weight could be a universal property of matter. For a long time, this observation underpinned the geocentric worldview powered by Aristoteles; he beliefed that heavier objects experience a higher gravitational acceleration.

Centuries later, in his famous book 'De revolutionibus orbium coelestium', Nicolaus Copernicus (1543) interpreted weight as divine phenomenon by which all things, including stars, planets and moons, are brought toward one another. In the 'Astronomia nova', Johannes Kepler (1609) compared weight with magnetism and hypothesized that any two stones attract each other in a way that is proportional to their masses. In the meantime, Galileo Galilei (1590) discovered that the acceleration of free falling test bodies at a given location does not depend on their masses, physical state or chemical composition. Modern measurements [8] confirm Galilei's discovery with a precision of a trillionth. In a vacuum, indeed, a one gram light feather and a one kilogram heavy lead ball experience the same acceleration of free fall. Long time before Friedrich Bessel (1832) and Lorand Eötvös (1908), Galileo Galilei's discovery was experimentally confirmed by Isaak Newton (1680) comparing the periods of pendulums of different masses but identical length. Nevertheless, in his universal law of gravitation, Newton (1687) postulated that gravity depends on the masses of the involved bodies. Though, he was deeply uncomfortable with this idea. 26 years after the first publication of his "Principia', in the age of 71, Newton wrote: "I have not yet been able to discover the cause of these properties of gravity from phenomena and I feign no hypotheses." Newton recognized the importance of not confusing gravity acceleration with the force that gravity can cause [9]. Actually, the question is not, does the force caused by gravity depend on the masses of the moving bodies. The question is rather, does mass *cause* the acceleration of free fall.

Analyzing the astronomical observations of Tycho Brahe, Johannes Kepler (1619) discovered that for every planet, the ratio of the cube of the semi-major axis R of the orbit and the square of the orbital period T is constant for a given orbital system. In the case of the Earth, this ratio defines the geocentric gravitational constant μ. Kepler's discovery is confirmed by high accuracy radar and laser ranging of the motion of artificial satellites. Thanks to Kepler's discovery, Earth's surface gravity acceleration can be derived from the orbital elements of any satellite, also from Moon's orbit:

$$g = \frac{\mu}{r^2} = \frac{\mu}{(6378000\ \text{m})^2} = 9.81\ \text{m/s}^2,$$

$$\mu = 4\pi^2 \frac{R^3}{T^2} = 3.9860044 \cdot 10^{14}\ \text{m}^3/\text{s}^2,$$

where R is the semi-major axis of Moon's orbit, T is the orbital period of the Moon and r is the equatorial radius of the Earth. No data about the masses or the chemical composition of the Earth or the Moon is needed.

Here it is important to underline that R and T are measured, but the identity $\mu = GM$ being the core of Newton's law of universal gravitation, is a theoretical presumption that provides mass as a source of gravity and the universality of the coefficient G as "gravitational constant'.

One of the basic principles of scientific research is the falsifiability of a theory. Obviously, any theory that postulates gravitation of mass as forming factor of the solar system is not falsifiable, because there is no method to *measure* the mass of a planet. Actually, no mass of any planet, planetoid or moon is measured, but only calculated based on the theoretical presumption $\mu = GM$.

Naturally, G is estimated in laboratory scale where masses can be measured. However, not only the correctness of the original experimental setup performed by Henry Cavendish (1798) is still under discussion, but also the correctness of more recent variants. There are large uncertainties not only in the obtained values of G, but even regarding the suitability of the applied methods of measuring gravity.

It is believed that gravitation cannot be screened. Because of this, it is virtually impossible to isolate the gravitational interaction between two masses from the presumed perturbative effects created by surrounding mass distributions. Invented by John Michell (1783), the instrument of choice for measuring G, the torsion pendulum, is subject to a variety of parasitic couplings and systematic effects which ultimately limit its suitability as a gravity transducer. George Gillies [10] listed about 350 papers almost all of which referred to work carried out with a torsion balance. Other sensitive mechanical devices are also pressed to the limits of their performance capabilities when employed for this purpose.

Besides of all the difficulties to measure G in laboratory, isn't there any other way to evidence the dependency of gravity on mass? For example, the Earth's surface masses are not uniformly distributed. There are huge mountains with a rock density of about three tons per cubic meter. There are oceans in which the density of water is only one ton per cubic meter - even at a depth of 10 kilometers. According to the logic of Newton's law of universal gravitation, these mass distribution inhomogeneities should act on sensitive gravimetric instruments. However, they do not [11].

In order to explain the absence of gravimetric evidence, the idea of isostasy [12] was invented. According to this hypothesis, the deeper the ocean, the more powerful the dense compensating deposits under its bottom; the higher the mountains, the looser is their foundation. Isostasy allegedly forms over huge periods of time, comparable to geological eras.

However, there are cases when very strong redistribution of surface masses occurs in a time period that is negligible by geological standards. For example, this happens during the eruption of an underwater volcano, when a seamount or even a new island builds up in a few days [13]. In these cases, there is no time to establish isostasy, and gravimetric instruments should react to these changes. Obviously, they do not react as expected, and for making gravity calculations more realistic, ground deformation data and numerical modelling is applied.

Gravimetric practice evidences that it is nearly impossible separating variations in gravity acceleration from low frequency seismic activity. Actually, gravimeters *are* long-period seismometers [14]. This is why the distribution of gravitational anomalies on gravity maps is indistinguishable from the zones of earthquakes and seismic activity.

Customarily, gravimetric data are recalculated with special corrections that providently consider the alleged effect of surface mass inhomogeneities. The corrections depend on the adopted model of the distribution of surface masses mainly based on seismic exploration. The idea to apply those corrections was proposed by Pierre Bouguer (1749). Now the difference between the really measured values of gravity and the theoretically calculated for an assumed mass density, is traditionally called a Bouguer-anomaly. Fluctuations in altitude of orbiting satellites indicating gravity variations are interpreted as caused by mass inhomogeneities [15]. In this way, gravimetric maps of planets and asteroids are being compiled.

In the case of mass as source of gravity, in accordance with Newton's shell theorem, a solid body with a spherically symmetric mass distribution should attract particles outside it as if its total mass were concentrated at its center. In contrast, the attraction exerted on a particle should decrease as the particle goes deeper into the body and it should become zero at the body's center.

The Preliminary Reference Earth Model [16] affirms the decrease of the gravity acceleration with the depth. However, this hypothesis is still under discussion. In 1981, Stacey, Tuck, Holding, Maher and Morris [17, 18] reported anomalous measures (larger values than expected) of the gravity acceleration in deep mines and boreholes. In [19] Frank Stacey writes that "geophysical measurements indicate a 1% difference between values at 10 cm and 1 km (depth); if confirmed, this observation will open up a new range of physics."

Furthermore, measurements of G are notoriously unreliable, so the constant is in permanent flux and the official value is an average. If G is changing, then G could depend on a new field. But this could also evidence that gravity itself may be changing. As mentioned Terry Quinn [20] of the Bureau International des Poids et Mesures (BIPM), the discrepant results may demonstrate that we do not understand the metrology of measuring weak force or signify some new physics.

Introduced with the postulated equation $\mu = GM$ as coefficient compensating the dimension of mass, G has no known confirmed dependence on any other fundamental constant. Suppose G would be estimated to be two times larger than the currently recommended value, this would simply mean that the masses of celestial bodies would be estimated to be two times smaller. However, this change would not have any impact on calculations depending on μ. In this case, the hypothesis that mass *causes* gravity, could turn out to be a dispensable assumption.

In view of this situation, it is understandable to intensify the search of possible derivations of G from theory. As mentioned Gillies [10], some recent approaches seek the ad hoc introduction of a new field or effect to create a situation in which a value for G can be built from ratios of other fundamental constants and numerical factors. However, most of the attempts come from a general relativistic starting point to examine the outcome of some scenario in which G arises from the calculations. For instance, Yanpeng Li [21] derives

$$G = \frac{1}{16\,\pi \cdot c \cdot \eta} = 6.636 \cdot 10^{-11}\ \mathrm{m^3 kg^{-1} s^{-2}}$$

from general relativity by introducing the "eigen-modulus of a tensor" as measure of its converging ability. According to Li, the eigen-modulus of the Einstein tensor equals 1/16 m/s^3, the mass density $\eta = 1$ kg/m^3 comes from the eigen-modulus of the energy-momentum tensor, c is the speed of light. Despite the numerical fit of the derived G value with the wide spectrum of data achieved in laboratory, the generality of this derivation and the physical sense of a mass density that equals 1 kg/m^3 may be questioned.

Introducing his geometric theory of gravitation, a century ago Einstein supposed that gravity is indistinguishable from, and in fact the same thing as, acceleration. Identifying gravity with acceleration $g = c \cdot f$, the gradient of a conservative gravitational field can be expressed in terms of frequency shifts:

$$\frac{\Delta f}{f} = g\,\frac{\Delta h}{c^2}\,.$$

Already in 1959, Robert Pound and Glen Rebka [22] verified this equation in their famous gravitational experiment. Sending gamma rays over a vertical distance of $\Delta h = 22.56$ m, they measured a blueshift of $\Delta f / f = 2.46 \cdot 10^{-15}$ that corresponds precisely with Earth's surface gravity 9.81 m/s^2.

Actually, also Kepler's 3rd law is of geometric origin and can be derived from Gauss's flux theorem in 3D-space within basic scale considerations. It applies to all conservative fields which decrease with the square of the distance, similar to the geometric dilution of the intensity of light into 3D-space.

The theoretical reduction of gravity to an acceleration enables the orbital motion to be identified with free fall. Orbital and rotational motions are periodic. So is free fall. Only the aggregate state of the planet prevents the free fall from becoming a damped oscillation. Considering gravity acting with the speed of light c, we can express gravity in units of time. For instance, Earth's surface gravity $g_{\mathrm{Earth}} = 9.81$ m/s^2 corresponds with an oscillation period of 355 days that is quite close to Earth's orbital period:

$$T_{\mathrm{Earth}} = \frac{c}{g_{\mathrm{Earth}}} = \frac{299792458\ \mathrm{m/s}}{9.81\ \mathrm{m/s^2}} = 355\,\mathrm{d}.$$

At an altitude of 100 km above sea level, Earth's gravity reduces down to 9.51 m/s^2 that corresponds with the orbital period of 365.25 days. In a series of experiments we demonstrated [23] that inside of finite spatial configurations which boundaries coincide with equipotential surfaces of the Fundamental Field (fig. 2), gravity acceleration reduces locally by 0.3 g down to 9.51 m/s^2.

The surface gravity $g_{\mathrm{Sun}} = 274$ m/s^2 of the Sun corresponds with an oscillation period of 12.7 days that is the first harmonic of its equatorial period 25.4 days of rotation. Similar coincidences are valid for the surface gravities of Mercury, Venus, Mars and even for Saturn and Jupiter. Although the definition of a planet's surface is conventional (especially in the case of gas giants), all these coincidences suggest the existence of an underlying connection of the gravity of a celestial body with its *own* orbital and rotational motions. Despite the rich history of crucial discoveries in astronomy and astrophysics and the development of sophisticated theories of gravitation, the distribution of stable orbits in the solar system remains to be little understood. In this context, the discovery of Johann Daniel Titius (1766) is even more remarkable. He found that the sequence of the planetary semi-major axes can be approximated by the exponential term:

$$a_n = 0.4 + 0.3 \cdot 2^n \,,$$

where the index n is $-\infty$ for Mercury, 0 for Venus, 1 for the Earth, 2 for Mars etc. Based on this idea, Johann Elert Bode, in 1772, first suggested that an undiscovered planet could exist between the orbits of Mars and Jupiter. William Herschel's discovery of Uranus in 1781 near the predicted distance 19.6 AU for the next body beyond Saturn increased faith in the law of Titius and Bode. In 1801, near the predicted for $n = 3$ distance 2.8 AU from Sun, Giuseppe Piazzi discovered the planetoid Ceres and the Franz Xaver von Zach group found further large asteroids.

In 1968, Stanley Dermott [24] found a similar progression for the major satellites of Jupiter, Saturn and Uranus. Nevertheless, at last, the hypothesis of Titius and Bode was discarded after it failed as a predictor of Neptune's orbit.

Surprisingly, recent astronomical research [25] suggests that exoplanetary systems follow Titius-Bode-like laws. Raw statistics from exoplanetary orbits indicate the exponential increase of semi-major axes as function of planetary index. It has been shown [2] that many exoplanetary systems follow an exponential progression of the form

$$a_n = a_0 + e^{bn}$$

with $n = 0, 1, 2, \ldots$; a_0 and b are constants to be determined for each system. Since its formulation, the Titius-Bode law has proved to be highly predictive, although its physical origin remains largely unclear.

Not only the distribution of stable orbits, but also the origin of the configuration of gravity fields in the solar system remains disputed. Furthermore, there is no known law concerning the rotation of celestial bodies besides conservation of the angular momentum [26] that they retain from the protoplanetary disks, so that the final distribution of the rotational periods appears as to be accidental.

In this article we demonstrate that the rotational and orbital periods of the planets, planetoids and large moons of the solar system as well as their gravitational constants approximate numeric attractors corresponding with the transcendental frequency ratios of scale-invariant eigenstates in chain systems of oscillating protons and electrons. The claims of our model we verify also on orbital periods of exoplanets and the gravitational constants of their stars.

Methods

In [27] we have shown that the difference between rational, irrational algebraic and transcendental numbers is not only a mathematical task, but it is also an essential aspect of stability in complex dynamic systems. For instance, integer frequency ratios provide resonance interaction that can destabilize a system [28]. Actually, it is transcendental numbers that define the preferred ratios of quantities which avoid destabilizing resonance interaction [29]. In this way, transcendental ratios of quantities sustain the lasting stability of periodic processes in complex dynamic systems. With reference to the evolution of a planetary system and its stability, we may therefore expect that the ratio of any two orbital periods should finally approximate a transcendental number.

Among all transcendental numbers, Euler's number $e = 2.71828\ldots$ is unique, because its real power function e^x coincides with its own derivatives. In the consequence, Euler's number allows inhibiting resonance interaction regarding any interacting periodic processes and their derivatives. Because of this unique property of Euler's number, complex dynamic systems tend to establish relations of quantities that coincide with values of the natural exponential function e^x for integer and rational exponents x.

Therefore, we expect that periodic processes in real systems prefer frequency ratios close to Euler's number and its rational powers. Consequently, the logarithms of their frequency ratios should be close to integer $0, \pm1, \pm2, \ldots$ or rational values $\pm1/2, \pm1/3, \pm1/4, \ldots$ In [30] we exemplified our hypothesis in particle physics, astrophysics, cosmology, geophysics, biophysics and engineering.

Based on this hypothesis, we introduced a fractal model of matter [31] as a chain system of harmonic quantum oscillators and could show the evidence of this model for all known hadrons, mesons, leptons and bosons as well. In [32] we have shown that the set of stable eigenstates in such systems is fractal and can be described by finite continued fractions:

$$\mathcal{F}_{jk} = \ln\left(\omega_{jk}/\omega_{00}\right) = \langle n_{j0}; n_{j1}, n_{j2}, \ldots, n_{jk}\rangle, \qquad (1)$$

where ω_{jk} is the set of angular eigenfrequencies and ω_{00} is the fundamental frequency of the set. The denominators are integer: $n_{j0}, n_{j1}, n_{j2}, \ldots, n_{jk} \in \mathrm{Z}$. The cardinality $j \in \mathrm{N}$ of the set and the number $k \in \mathrm{N}$ of layers are finite. In the canonical form, all numerators equal 1. We use angle brackets for continued fractions.

Any finite continued fraction represents a rational number [33]. Therefore, the ratios ω_{jk}/ω_{00} of eigenfrequencies are always irrational, because for rational exponents the natural exponential function is transcendental [34]. This circumstance provides for lasting stability of those eigenstates of a chain system of harmonic oscillators because it prevents resonance interaction [35] between the elements of the system.

The distribution density of stable eigenstates reaches local maxima near reciprocal integers $\pm1/2, \pm1/3, \pm1/4, \ldots$ that are attractor points (fig. 1) in the fractal set $\mathcal{F}_{jk}$ of natural logarithms. Integer logarithms $0, \pm1, \pm2, \ldots$ represent the most stable eigenstates (main attractors).

In the case of harmonic quantum oscillators, the continued fractions $\mathcal{F}_{jk}$ define not only fractal sets of natural angular frequencies ω_{jk}, angular accelerations $a_{jk} = c \cdot \omega_{jk}$, oscillation periods $\tau_{jk} = 1/\omega_{jk}$ and wavelengths $\lambda_{jk} = c/\omega_{jk}$ of the chain system, but also fractal sets of energies $\mathrm{E}_{jk} = \hbar \cdot \omega_{jk}$ and masses $\mathrm{m}_{jk} = \mathrm{E}_{jk}/c^2$ which correspond with the eigenstates of the system. For this reason, we call the continued fraction $\mathcal{F}_{jk}$ the *Fundamental Fractal* of stable eigenstates in chain systems of harmonic quantum oscillators.

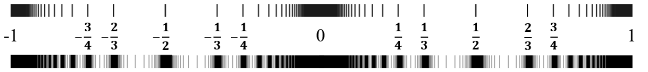


Fig. 1: The distribution of stable eigenvalues of $\mathcal{F}_{jk}$ for $k = 1$ (above) and for $k = 2$ (below) in the range $-1 \leqslant \mathcal{F}_{jk} \leqslant 1$.

The spatio-temporal projection of the Fundamental Fractal $\mathcal{F}_{jk}$ of stable eigenstates is a fractal scalar field of transcendental attractors, the *Fundamental Field* [36].

The connection between the spatial and temporal projections of the Fundamental Fractal is given by the speed of light c = 299792458 m/s. The constancy of c makes both projections isomorphic, so that there is no arithmetic or geometric difference. Only the units of measurement are different.

Figure 2 shows the linear 2D-projection $\exp(\mathcal{F}_{jk})$ of the first layer of the Fundamental Field

$$\mathcal{F}_{\mathrm{j}1} = \langle n_{j0}; n_{j1} \rangle = n_{j0} + \frac{1}{n_{j1}}$$

in the interval $-1 < \mathcal{F}_{\mathrm{j}1} < 1$. The upper part of figure 1 shows the same interval in the logarithmic representation. The Fundamental Field is topologically 3-dimensional, a fractal set of embedded spheric equipotential surfaces. The logarithmic potential difference defines a gradient directed to the center of the field that causes a central force of attraction. Because of the fractal logarithmic hyperbolic metric of the field, every equipotential surface is an attractor. The scalar potential difference $\Delta\mathcal{F}$ of sequent equipotential surfaces at a given layer k is defined by the difference of continued fractions (1):

$$\Delta\mathcal{F} = \mathcal{F}(\mathrm{j,k}) - \mathcal{F}(\mathrm{j+1,k}) =$$
$$= \langle n_{j0}; n_{j1}, n_{j2}, \ldots, n_{jk} \rangle - \langle n_{j0}; n_{j1}, n_{j2}, \ldots, n_{j+1,k} \rangle .$$

For instance, at the first layer k=1, the potential differences have the form:

$$\Delta\mathcal{F} = \frac{1}{n_{j1}} - \frac{1}{n_{j1}+1} = \frac{1}{n_{j1}^2 + n_{j1}} .$$

Therefore, the potential difference between sequent equipotential surfaces at any given layer $k+1$ decreases parabolically, approximating zero near an equipotential surface of the layer k. This is why any equipotential surface is an attractor where potential differences decrease and processes can gain stability. Main attractors at the layer $k = 0$ correspond with integer logarithms, subattractors at deeper layers $k > 0$ correspond with rational logarithms.

The Fundamental Field is of pure arithmetical origin, and there is no particular physical mechanism required as field source. It is all about transcendental ratios of frequencies [29] that inhibit destabilizing resonance. In this way, the Fundamental Field concerns all repetitive processes which share at least one characteristic — the frequency. Therefore, we postulate the universality of the Fundamental Field that affects any type of physical interaction, regardless of its complexity.

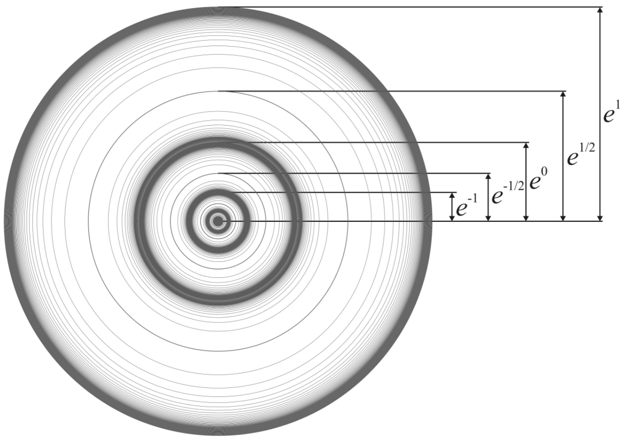


Fig. 2: The equipotential surfaces of the Fundamental Field in the linear 2D-projection for k = 1.

In fact, scale relations in particle physics [31, 37, 38], nuclear physics [39, 40] and astrophysics [4] obey the same Fundamental Fractal (1), without any additional or particular settings. The proton-to-electron rest energy ratio approximates the first layer of the Fundamental Fractal that could explain their exceptional stability [30]. The life-spans of the proton and electron top everything that is measurable, exceeding 10^{29} years [41].

Property	Electron	Proton
$\mathrm{E} = \mathrm{mc}^2$	0.5109989461(31) MeV	938.2720813(58) MeV
$\omega = \mathrm{E}/\hbar$	$7.76344 \cdot 10^{20}$ Hz	$1.42549 \cdot 10^{24}$ Hz
$\tau = 1/\omega$	$1.28809 \cdot 10^{-21}$ s	$7.01515 \cdot 10^{-25}$ s
$\lambda = c/\omega$	$3.86159 \cdot 10^{-13}$ m	$2.10309 \cdot 10^{-16}$ m

Table 1: The basic set of the physical properties of the electron and proton. Data from Particle Data Group [41]. Frequencies, oscillation periods and wavelengths are calculated.

These unique properties of the electron and proton predestinate their physical characteristics as fundamental units. Table 1 shows the basic set of electron and proton units that can be considered as a fundamental metrology (c is the speed of light in a vacuum, $\hbar$ is the Planck constant). In [32] was shown that the fundamental metrology (tab. 1) is completely compatible with Planck units [42]. Originally proposed in 1899 by Max Planck, these units are also known as natural units, because the origin of their definition comes only from properties of nature and not from any human construct. Max Planck wrote [43] that these units, "regardless of any particular bodies or substances, retain their importance for all times and for all cultures, including alien and non-human, and can therefore be called natural units of measurement". Planck units reflect the characteristics of space-time.

We hypothesize that scale invariance according the Fundamental Fractal (1) calibrated on the physical properties of the proton and electron is a universal characteristic of organized matter and criterion of stability. This hypothesis we have called *Global Scaling* [30].

On this background, atoms and molecules emerge as stable eigenstates in fractal chain systems of harmonically oscillating protons and electrons. Andreas Ries [38] demonstrated that this model allows for the prediction of the most abundant isotope of a given chemical element.

In [44] we applied the Fundamental Fractal (1) to macroscopic scales interpreting gravity as attractor effect of its stable eigenstates. Indeed, the orbital and rotational periods of planets, planetoids and large moons of the solar system correspond with attractors of electron and proton stability [32]. This is valid also for the planets [30] of the systems Trappist 1 and Kepler 20. Planetary and lunar orbits [4] correspond with equipotential surfaces of the Fundamental Field, as well as the metric characteristics of stratification layers in planetary atmospheres [45]. In [36] we demonstrated that the Fundamental Field (fig. 2) in the interval of the main attractors $\langle 49\rangle \leqslant \mathcal{F} \leqslant \langle 52\rangle$ of proton stability reproduces the 2D profile of the Earth's interior confirmed by seismic exploration.

Results

We will show now that the orbital and rotational periods of planets, planetoids and moons as well as their gravity accelerations approximate stable eigenstates of our model of matter as fractal chain system of oscillating protons and electrons, described by the Fundamental Fractal.

In accordance with the equation (1), we calculate the natural logarithm of the ratio of the measured value to the corresponding electron or proton unit taken from table 1. For instance, the orbital period of Jupiter T_{O} (Jupiter) = 4332.59 days = $3.7434 \cdot 10^8$ seconds [46] matches the main attractor $\mathcal{F}\langle 66\rangle$ of *electron* stability:

$$\ln\left(\frac{T_{\mathrm{O}}(\text{Jupiter})}{2\pi \cdot \tau_e}\right) = \ln\left(\frac{3.7434 \cdot 10^8\,\mathrm{s}}{2\pi \cdot 1.28809 \cdot 10^{-21}\,\mathrm{s}}\right) = 66.00\,.$$

In contrast to orbital motion, rotation is an angular motion, so that the proton or electron angular oscillation periods are applied as units. The rotation period T_{R}(Ceres) = 9 hours = 32400 seconds of Ceres, the largest body of the main asteroid belt, matches the main attractor $\mathcal{F}\langle 66\rangle$ of *proton* stability:

$$\ln\left(\frac{T_{\mathrm{R}}(\text{Ceres})}{\tau_p}\right) = \ln\left(\frac{32400\,\mathrm{s}}{7.01515 \cdot 10^{-25}\,\mathrm{s}}\right) = 66.00\,.$$

Table 3 gives an overview of the orbital and rotational periods as well as the gravitational constants of the planets including the planetoid Ceres and large moons.

Within our model, the approximation level of an attractor of stability indicates evolutionary trends. For instance, Venus' OE2 = 63.04 indicates that the orbital period of the Morning star must slightly decrease for reaching the center of the main attractor $\mathcal{F}\langle 63\rangle$. On the contrary, Moon's OE2 = 60.94 indicates that its orbital period must still increase for reaching the center of the main attractor $\mathcal{F}\langle 61\rangle$. Actually, exactly this is observed [47]. As well, Uranus' OE2 = 67.96 let us expect an increase of its orbital period in order to reach the main attractor $\mathcal{F}\langle 68\rangle$. Mercury's OE1 = 63.94 indicates that in future it could overcome the current tidal 3/2 locking by reaching the main attractor $\mathcal{F}\langle 64\rangle$ of electron stability. Mercury's RP1 = 71.05 indicates that its rotation must speed up slightly [26] in order to reach the attractor $\mathcal{F}\langle 71\rangle$ of proton stability. Earth's RP1 = 66.98 indicates that our planet must slow its rotation by 24 minutes per turn in order to reach the main attractor $\mathcal{F}\langle 67\rangle$.

Despite conservation of angular momentum [26], there is no known law concerning the rotation of celestial bodies. The more remarkable is the correspondence of the rotation periods of planets, planetoids and large moons with attractors of the Fundamental Fractal (1) as shown in table 3.

For instance, Mars, Ceres and Jupiter have reached the main attractor $\mathcal{F}\langle 66\rangle$ in quite different way. In the case of Mars and Jupiter, the attractor $\mathcal{F}\langle 66\rangle$ stabilizes the orbital period T_{O}. In the case of the planetoid Ceres, the same attractor $\mathcal{F}\langle 66\rangle$ stabilizes the period of rotation T_{R}. Actually, the difference lays in the reference units. In the case of Jupiter's orbital period, the reference unit is the oscillation period of the electron $2\pi\tau_e$; in the case of Mars, it is the angular oscillation period of the electron τ_e, and in the case of the rotational period of Ceres, it is the angular oscillation period of the proton τ_p. Now we can write down the following relations:

$$T_{\mathrm{O}}(\text{Jupiter}) = 2\pi \cdot T_{\mathrm{O}}(\text{Mars}),$$

$$T_{\mathrm{O}}(\text{Mars}) = \frac{\tau_e}{\tau_p} \cdot T_{\mathrm{R}}(\text{Ceres}).$$

The complete (polar) rotational period of the Sun approximates the main attractor $\mathcal{F}\langle 63\rangle$ of electron stability:

$$\ln\left(\frac{T_{\mathrm{R}}(\text{Sun})}{\tau_e}\right) = 63.01\,.$$

The orbital period of Venus approximates the same attractor $\mathcal{F}\langle 63\rangle$, as table 3 shows. Consequently, the scaling factor 2π

connects the orbital period of Venus with the rotational period of the Sun:

$$T_O(\text{Venus}) = 2\pi \cdot T_R(\text{Sun}).$$

Archimedes' number $\pi = 3.14159\ldots$ is transcendental and therefore, it does not violate the principle of avoiding destabilizing resonance. Needless to say that these relations cannot be derived from Kepler's laws or Newton's law of gravitation. The proton-to-electron ratio (tab. 1) approximates the seventh power of Euler's number and its square root:

$$\ln\left(\frac{\omega_p}{\omega_e}\right) = \ln\left(\frac{1.42549 \cdot 10^{24}\,\text{Hz}}{7.76344 \cdot 10^{20}\,\text{Hz}}\right) \simeq 7 + \frac{1}{2} = \langle 7; 2 \rangle.$$

In the consequence of this potential difference of the proton relative to the electron, the scaling factor $\sqrt{e} = 1.64872\ldots$ connects attractors of proton stability with similar attractors of electron stability in alternating sequence. The following Diophantine equation describes the correspondence of proton calibrated attractors n_p with electron calibrated attractors n_e. Non considering the signature, only three pairs (n_p, n_e) of integers are solutions to this equation: (3, 6), (4, 4), (6, 3).

$$\frac{1}{n_p} + \frac{1}{n_e} = \frac{1}{2}\,.$$

Figure 3 demonstrates this situation on the first layer of the Fundamental Fractal (1). Both, the attractors of proton and electron stability are represented at the first layer, so we can see clearly that among the integer or half, only the attractors ±1/3, ±1/4 and ±1/6 are common. In these attractors, proton stability is supported by electron stability and vice versa, so we expect that they are preferred in real systems.

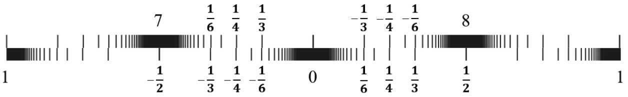


Fig. 3: The distribution of the attractors of proton (bottom) stability in the range $-1 < \mathcal{F} < 1$ of the attractors of electron (top) stability. Natural logarithmic representation.

Figure 4 shows the distribution of the number of exoplanets with orbital periods in the range 5 d $< T_O <$ 24 d that corresponds with the range of logarithms $59.2 < \ln(T_O/2\pi\tau_e) < 60.8$ on the horizontal axis. According with table 1, τ_e is the electron angular oscillation period. The histogram contains data of 1430 exoplanets and shows clearly the maximum corresponding with the main attractor $\mathcal{F}\langle 60 \rangle$. Other maxima correspond with the attractors $\mathcal{F}\langle 59; 2 \rangle$ and $\mathcal{F}\langle 60; 2 \rangle$; even the subattractors $\mathcal{F}\langle 60; -4 \rangle$ and $\mathcal{F}\langle 60; 4 \rangle$ can be distinguished.

The histogram evidences that the majority of the analyzed 1430 exoplanets [48] prefer orbital periods close to 10–11 days corresponding with the main attractor $\mathcal{F}\langle 60 \rangle$, as well as periods close to 6–7 days or close to 17–18 days corresponding with the attractors $\mathcal{F}\langle 59; 2 \rangle$ and $\mathcal{F}\langle 60; 2 \rangle$. Because of the logarithm 7+1/2 of the proton-to-electron ratio, the attractors $\mathcal{F}\langle 59; 2 \rangle$ and $\mathcal{F}\langle 60; 2 \rangle$ of electron stability are actually the main attractors $\mathcal{F}\langle 67 \rangle$ and $\mathcal{F}\langle 68 \rangle$ of proton stability.

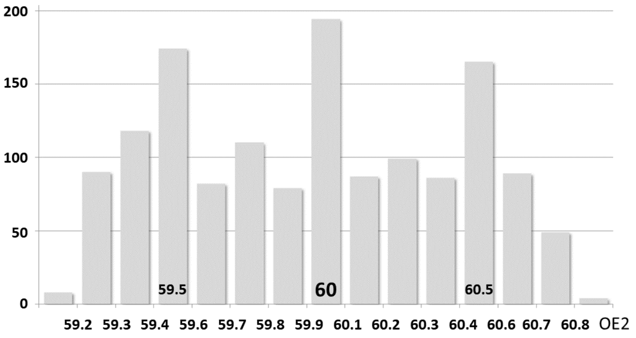


Fig. 4: The histogram shows the distribution of the number of exoplanets with orbital periods in the range 5 d $< T_O <$ 24 d. The logarithms OE2 = $\ln(T_O/2\pi\tau_e)$ are on the horizontal axis. Corresponding with table 1, τ_e is the electron angular oscillation period. Data of 1430 exoplanets taken from [48].

Now we can also explain the origin of the Titius-Bode law. The OE2 column in tab. 3 shows that the orbital periods of Ceres, Jupiter, Saturn and Uranus approximate the sequence of the main attractors $\mathcal{F} = \langle 65 \rangle$, $\langle 66 \rangle$, $\langle 67 \rangle$ and $\langle 68 \rangle$ of electron stability. The ratio of main attractors equals Euler's number $e = 2.71828\ldots$ Considering Kepler's third law, from this directly follows that the ratio of the semi-major axes of Ceres, Jupiter, Saturn and Uranus approximates the cube root of the square of Euler's number $e^{2/3} = 1.9477\ldots$ This is why the Titius-Bode law approximates the exponential function 2^n. However, not all orbital periods approximate main attractors. The Earth-Venus orbital period ratio approximates the square root of Euler's number. Consequently, the ratio of their semi-major axes approximates the cube root of Euler's number $e^{1/3} = 1.3956\ldots$ The same is valid for Umbriel and Ariel, the moons of Uranus. The Neptune-Uranus orbital period ratio approximates $e^{2/3}$. Consequently, the ratio of their semi-major axes approximates $e^{4/9} = 1.5596\ldots$

The eigenvalues of $\mathcal{F}$ are transcendental, and their distribution (1) is logarithmically fractal. This is why Titius-Bode-like equations cannot deliver a general and complete model of an orbital system.

Among the orbital and rotational periods, tab. 3 shows that also the gravitational constants μ obey the Fundamental Fractal (1) approximating main attractors and the preferred subattractors as shown in fig. 3.

In accordance with [46], surface gravities g are given for a distance from the center of the celestial body that coincides with the radius of the solid or liquid surface, without consideration of the centrifugal effects of rotation. For gas giants such as Jupiter, Saturn, Uranus, and Neptune, where the surfaces are deep in the atmosphere and the radius is not known, the surface gravity is given at the 1 bar pressure level in the atmosphere. In this way, any surface gravity is given for an individual distance from the local center of gravitation.

Earth's surface gravity corresponds to the equatorial radius at sea level 6378 km, and the surface gravity of Uranus corresponds to its equatorial radius of 25559 km where the atmospheric pressure equals 1 bar. Although the surface gravities on Venus and Uranus are identical equal 8.87 m/s^2, this does not mean that they indicate comparable gravitational fields. Therefore, we cannot use the surface gravity accelerations for comparison, but only the gravitational constants μ.

Star	μ, m^3/s^2	MP	$\mathcal{F}$	MP - $\mathcal{F}$
Trappist 1	$1.1976 \cdot 10^{19}$	40.99	$\langle 41 \rangle$	-0.01
Proxima Cent	$1.5725 \cdot 10^{19}$	41.26	$\langle 41; 4 \rangle$	0.01
Gliese 1061	$1.6966 \cdot 10^{19}$	41.34	$\langle 41; 3 \rangle$	0.01
Barnard's star	$2.6154 \cdot 10^{19}$	41.77	$\langle 42; -4 \rangle$	0.02
Struve 2398 B	$3.7765 \cdot 10^{19}$	42.14	$\langle 42; 6 \rangle$	-0.02
Gliese 876	$4.2851 \cdot 10^{19}$	42.27	$\langle 42; 4 \rangle$	0.02
Lacaille 9352	$6.4378 \cdot 10^{19}$	42.67	$\langle 43; -3 \rangle$	0.00
Tau Ceti	$1.0414 \cdot 10^{20}$	43.15	$\langle 43; 6 \rangle$	-0.01
HD 69830	$1.1402 \cdot 10^{20}$	43.24	$\langle 43; 4 \rangle$	-0.01
55 Cancri	$1.2480 \cdot 10^{20}$	43.33	$\langle 43; 3 \rangle$	0.00
Upsilon Andro	$1.7598 \cdot 10^{20}$	43.68	$\langle 44; -3 \rangle$	0.01

Table 2: The gravitational constants μ of some stars calculated from data [48] of orbital periods and semi-major axes of their planets. MP = ln $(\mu/\lambda_p^3\omega_p^2)$. Corresponding with tab. 1, λ_p is the proton angular wavelength and ω_p is the proton angular frequency. Continued fractions (1) of the Fundamental Fractal $\mathcal{F}$ are given in angle brackets.

Table 3 shows that the gravitational constants μ of Pluto, Neptune, Jupiter, Mars and Venus approximate main attractors $\mathcal{F} = \langle n_0 \rangle$ of electron stability. The gravitational constants of the other planets and planetoids of the solar system approximate the rational subattractors $\mathcal{F} = \langle n_0 \pm 1/2 \rangle$, $\langle n_0 \pm 1/3 \rangle$, $\langle n_0 \pm 1/4 \rangle$ or $\langle n_0 \pm 1/6 \rangle$. As well, the gravitational constants of the large moons of Jupiter, Saturn, Uranus and Neptune approximate main attractors of electron and proton stability and the same rational subattractors. This is valid also for exoplanetary systems. Table 2 shows the gravitational constants μ of some near stars calculated from data [48] of the orbital periods and semi-major axes of their planets.

Conclusion

Perhaps, the conventional paradigm of physical interaction should be completed by the principle of avoiding those interactions that potentially can destabilize a system.

Admittedly, the principle of minimum action is an essential part of theoretical physics at least since Pierre de Fermat (1662) and Pierre Louis Moreau de Maupertuis (1741), Joseph-Louis Lagrange (1788) and William Rowan Hamilton (1834) applied in the Euler – Lagrange equations of motion.

The novelty of our solution we see in the purely numerical approach that rediscovers Euler's number, its integer powers and roots as attractors of transcendental numbers. Approximating transcendental ratios of quantities defined by integer and rational natural logarithms, complex dynamic systems can avoid destabilizing resonance interactions between their elements and gain lasting stability. As we have shown in this paper, planetary systems make extensive use of this solution.

Finally, we can explain why Jupiter's orbital period equals 4332.59 days: With this orbital period, Jupiter occupies the main equipotential surface $\mathcal{F} = \langle 66 \rangle$ of the Fundamental Field of transcendental attractors and in this way, Jupiter avoids destabilizing resonance interactions with the orbital motions of other planets and gains lasting stability of its own orbital motion. In other words, there is a fractal scalar field of transcendental temporal attractors corresponding with integer and rational powers of Euler's number. One of these attractors is $\mathcal{F} = \langle 66 \rangle$, and it has materialized as a stable orbital period in the solar system among the attractors $\mathcal{F} = \langle 62 \rangle$, $\langle 63 \rangle$, $\langle 64 \rangle$, $\langle 65 \rangle$, $\langle 67 \rangle$, $\langle 68 \rangle$, $\langle 69 \rangle$ and their subattractors. Smaller attractors $\mathcal{F} = \langle 58 \rangle$, $\langle 59 \rangle$, $\langle 60 \rangle$ and $\langle 61 \rangle$ and their subattractors define stable orbital periods in moon systems and in the majority of the discovered so far exoplanetary systems.

Naturally, the Fundamental Field $\mathcal{F}$ of transcendental attractors does not materialize in the scale of planetary systems only. At subatomic scale, it defines the proton-to-electron ratio and in this way, allows the formation of stable atoms and complex matter. At planetary scale, now we can distinguish attractors of electron stability and attractors of proton stability. While the attractors of electron stability define stable orbital periods, the attractors of proton stability define stable rotational periods. For instance, the attractor $\mathcal{F} = \langle 66 \rangle$ of *electron* stability defines the orbital period of Jupiter, and the same attractor $\mathcal{F} = \langle 66 \rangle$ of *proton* stability defines the rotational period of Mars. In this way, the law behind the distribution of stable orbital and rotational periods is the same Fundamental Field of transcendental attractors.

Interpreting gravity in terms of frequency, we did demonstrate that the distribution of gravity in the solar system is not accidental, but obeys the same Fundamental Field $\mathcal{F}$. As well, the gravitational constants μ of extrasolar systems obey the logarithmically fractal metric (1) of $\mathcal{F}$. This circumstance let us suppose that even entire planetary systems prefer avoiding destabilizing resonance interactions between them.

Acknowledgements

The author is grateful to Simon Shnoll, Viktor Panchelyuga, Valery Kolombet, Oleg Kalinin, Viktor Bart, Andreas Ries, Michael Kauderer, Ulrike Granögger and Leili Khosravi for valuable discussions.

Submitted on January 7, 2021

Body	T_O, d	OE1	$\mathcal{F}$	OE2	$\mathcal{F}$	T_R, h	RP1	$\mathcal{F}$	RP2	$\mathcal{F}$	μ, m^3/s^2	ME	$\mathcal{F}$
Eris	204199.00	71.69	$\langle 72; -3 \rangle$	69.86	$\langle 70; -6 \rangle$	349.44	69.66	$\langle 70; -3 \rangle$	67.82	$\langle 68; -6 \rangle$	$1.10800 \cdot 10^{12}$	17.28	$\langle 17; 4 \rangle$
Pluto	90560.09	70.88	$\langle 71; -6 \rangle$	69.04	$\langle 69 \rangle$	153.29	68.84	$\langle 69; -6 \rangle$	67.00	$\langle 67 \rangle$	$8.62000 \cdot 10^{11}$	17.03	$\langle 17 \rangle$
Neptune	60193.20	70.47	$\langle 70; 2 \rangle$	68.64	$\langle 69; -3 \rangle$	16.11	66.58	$\langle 66; 2 \rangle$	64.75	$\langle 65; -4 \rangle$	$6.83653 \cdot 10^{15}$	26.01	$\langle 26 \rangle$
Uranus	30688.49	69.80	$\langle 70; -6 \rangle$	67.96	$\langle 68 \rangle$	17.24	66.65	$\langle 67; -3 \rangle$	64.81	$\langle 65; -6 \rangle$	$5.79394 \cdot 10^{15}$	25.84	$\langle 26; -6 \rangle$
Saturn	10759.21	68.75	$\langle 69; -4 \rangle$	66.91	$\langle 67 \rangle$	10.56	66.16	$\langle 66; 6 \rangle$	64.32	$\langle 64; 3 \rangle$	$3.79312 \cdot 10^{16}$	27.72	$\langle 28; -4 \rangle$
Jupiter	4332.60	67.84	$\langle 68; -6 \rangle$	66.00	$\langle 66 \rangle$	9.93	66.10	$\langle 66; 6 \rangle$	64.26	$\langle 64; 3 \rangle$	$1.26687 \cdot 10^{17}$	28.93	$\langle 29 \rangle$
Ceres	1683.80	66.90	$\langle 67; -6 \rangle$	65.06	$\langle 65 \rangle$	9.00	66.00	$\langle 66 \rangle$	64.16	$\langle 64; 6 \rangle$	$6.26274 \cdot 10^{10}$	14.41	$\langle 14; 2 \rangle$
Mars	686.97	66.00	$\langle 66 \rangle$	64.16	$\langle 64; 6 \rangle$	24.62	67.01	$\langle 67 \rangle$	65.17	$\langle 65; 6 \rangle$	$4.28284 \cdot 10^{13}$	20.93	$\langle 21 \rangle$
Earth	365.25	65.37	$\langle 65; 3 \rangle$	63.53	$\langle 63; 2 \rangle$	24.00	66.98	$\langle 67 \rangle$	65.15	$\langle 65; 6 \rangle$	$3.98600 \cdot 10^{14}$	23.16	$\langle 23; 6 \rangle$
Venus	224.70	64.88	$\langle 65; -6 \rangle$	63.04	$\langle 63 \rangle$	243.03	72.48	$\langle 72; 2 \rangle$	70.64	$\langle 71; -3 \rangle$	$3.24859 \cdot 10^{14}$	22.96	$\langle 23 \rangle$
Mercury	87.97	63.94	$\langle 64 \rangle$	62.11	$\langle 62; 6 \rangle$	58.65	71.05	$\langle 71 \rangle$	69.22	$\langle 69; 6 \rangle$	$2.20320 \cdot 10^{13}$	20.27	$\langle 20; 4 \rangle$
Moon	27.32	62.78	$\langle 63; -6 \rangle$	60.94	$\langle 61 \rangle$	sync	70.29	$\langle 70; 3 \rangle$	68.45	$\langle 68; 2 \rangle$	$4.90487 \cdot 10^{12}$	18.77	$\langle 19; -4 \rangle$
Callisto	16.69	62.28	$\langle 62; 3 \rangle$	60.44	$\langle 60; 2 \rangle$	sync	69.80	$\langle 70; -6 \rangle$	67.96	$\langle 68 \rangle$	$7.17929 \cdot 10^{12}$	19.15	$\langle 19; 6 \rangle$
Ganymede	7.15	61.44	$\langle 61; 2 \rangle$	59.60	$\langle 60; -3 \rangle$	sync	68.95	$\langle 69 \rangle$	67.11	$\langle 67; 6 \rangle$	$9.88783 \cdot 10^{12}$	19.47	$\langle 19; 2 \rangle$
Europa	3.55	60.74	$\langle 61; -4 \rangle$	58.90	$\langle 59 \rangle$	sync	68.25	$\langle 68; 4 \rangle$	66.41	$\langle 66; 2 \rangle$	$3.20274 \cdot 10^{12}$	18.34	$\langle 18; 3 \rangle$
Io	1.77	60.04	$\langle 60 \rangle$	58.20	$\langle 58; 6 \rangle$	sync	67.55	$\langle 67; 2 \rangle$	65.72	$\langle 66; -3 \rangle$	$5.95992 \cdot 10^{12}$	18.96	$\langle 19 \rangle$
Iapetus	79.32	63.84	$\langle 64; -6 \rangle$	62.00	$\langle 62 \rangle$	sync	71.36	$\langle 71; 3 \rangle$	69.52	$\langle 69; 2 \rangle$	$1.20500 \cdot 10^{11}$	15.06	$\langle 15 \rangle$
Titan	15.95	62.24	$\langle 62; 4 \rangle$	60.40	$\langle 60; 2 \rangle$	sync	69.75	$\langle 70; -4 \rangle$	67.91	$\langle 68 \rangle$	$8.96273 \cdot 10^{12}$	19.37	$\langle 19; 3 \rangle$
Rhea	4.52	60.98	$\langle 61 \rangle$	59.14	$\langle 59; 6 \rangle$	sync	68.49	$\langle 69; 2 \rangle$	66.65	$\langle 67; -3 \rangle$	$1.54000 \cdot 10^{11}$	15.31	$\langle 15; 3 \rangle$
Dione	2.74	60.47	$\langle 60; 2 \rangle$	58.64	$\langle 59; -3 \rangle$	sync	67.99	$\langle 68 \rangle$	66.15	$\langle 66; 6 \rangle$	$7.10000 \cdot 10^{10}$	14.53	$\langle 14; 2 \rangle$
Tethys	1.89	60.10	$\langle 60; 6 \rangle$	58.27	$\langle 58; 3 \rangle$	sync	67.62	$\langle 68; -3 \rangle$	65.78	$\langle 66; -6 \rangle$	$4.12000 \cdot 10^{10}$	13.99	$\langle 14 \rangle$
Enceladus	1.37	59.78	$\langle 60; -6 \rangle$	57.94	$\langle 58 \rangle$	sync	67.30	$\langle 67; 3 \rangle$	65.46	$\langle 65; 2 \rangle$	$7.20000 \cdot 10^{9}$	12.24	$\langle 12; 4 \rangle$
Mimas	0.94	59.41	$\langle 59; 3 \rangle$	57.57	$\langle 57; 2 \rangle$	sync	66.92	$\langle 67 \rangle$	65.09	$\langle 65 \rangle$	$2.50000 \cdot 10^{9}$	11.18	$\langle 11; 6 \rangle$
Oberon	13.46	62.07	$\langle 62 \rangle$	60.23	$\langle 60; 6 \rangle$	sync	69.58	$\langle 69; 2 \rangle$	67.75	$\langle 68; -4 \rangle$	$1.93000 \cdot 10^{11}$	15.53	$\langle 15; 2 \rangle$
Titania	8.71	61.63	$\langle 62; -3 \rangle$	59.79	$\langle 60; -6 \rangle$	sync	69.15	$\langle 69; 6 \rangle$	67.31	$\langle 67; 3 \rangle$	$2.20000 \cdot 10^{11}$	15.66	$\langle 16; -3 \rangle$
Umbriel	4.14	60.89	$\langle 61; -6 \rangle$	59.05	$\langle 59 \rangle$	sync	68.40	$\langle 68; 3 \rangle$	66.57	$\langle 66; 2 \rangle$	$8.95000 \cdot 10^{10}$	14.76	$\langle 15; -4 \rangle$
Ariel	2.52	60.39	$\langle 60; 3 \rangle$	58.55	$\langle 58; 2 \rangle$	sync	67.91	$\langle 68; -6 \rangle$	66.07	$\langle 66 \rangle$	$7.88000 \cdot 10^{10}$	14.64	$\langle 15; -3 \rangle$
Miranda	1.41	59.81	$\langle 60; -6 \rangle$	57.98	$\langle 58 \rangle$	sync	67.33	$\langle 67; 3 \rangle$	65.49	$\langle 65; 2 \rangle$	$4.00000 \cdot 10^{9}$	11.65	$\langle 12; -3 \rangle$
Triton	5.88	61.24	$\langle 61; 4 \rangle$	59.40	$\langle 59; 2 \rangle$	sync	68.75	$\langle 69; -4 \rangle$	66.92	$\langle 67 \rangle$	$1.42689 \cdot 10^{12}$	17.53	$\langle 17; 2 \rangle$

Table 3: The sidereal orbital periods T_O, rotational periods T_R and gravitational constants μ of the planets, planetoids and large moons of the solar system. OE1 = ln (T_O/τ_e), OE2 = ln $(T_O/2\pi\tau_e)$, RP1 = ln (T_R/τ_p), RP2 = ln $(T_R/2\pi\tau_p)$, ME = ln $(\mu/\lambda_p^3\omega_p^2)$. Corresponding with tab. 1, τ_e is the *electron* angular oscillation period, τ_p is the *proton* angular oscillation period, λ_e is the *electron* angular wavelength and ω_e is the *electron* angular frequency. The continued fractions (1) of the Fundamental Fractal $\mathcal{F}$ are given in angle brackets. Although some data is shown with two decimals only, for calculating the logarithms, high precision data [46, 49–51] were used.

References

1. Verlinde E. On the origin of gravity and the laws of Newton. *arXiv*:1001.0785v1 [hep-th] 6 Jan 2010.
2. Lara P., Cordero-Tercero G., Allen Ch. The reliability of the Titius-Bode relation and its implications for the search for exoplanets. *arXiv*:2003.05121v1 [astro-ph.EP] 11 Mar 2020.
3. Brookes C. J. On the prediction of Neptune. Provided by the NASA Astrophysics Data System. *Kluwer Academic Publishers*, (1970).
4. Müller H. Global Scaling of Planetary Systems. *Progress in Physics*, 2018, v. 14, 99–105.
5. Gillon M. et al. Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1. *Nature*, 21360, 2017.
6. Hand E. Kepler discovers first Earth-sized exoplanets. *Nature*, 9688, 2011.
7. Turyshev S. G., Toth V. T. The Puzzle of the Flyby Anomaly. *arXiv*:0907.4184v1 [gr-qc] 23 Jul 2009.
8. Schlamminger S. et al. Test of the Equivalence Principle Using a Rotating Torsion Balance. *arXiv*:0712.0607v1 [gr-qc] 4 Dec 2007.
9. Janiak A. Newton and the Reality of Force. *Journal of the History of Philosophy*, vol. 45, Nr. 1 (2007) 127–147.
10. Gillies G. T. The Newtonian gravitational constant: recent measurements and related studies. *Rep. Prog. Phys*, vol. 60, 151–225, (1997).
11. Stacey F. D., Tuck G. J. Geophysical evidence for non-newtonian gravity. *Nature*, v. 292, pp. 230–232, (1981).
12. Miller A. H. Gravity and Isostasy. *Journal of the Royal Astronomical Society of Canada*, vol. 20, p. 327, (1926).
13. Battaglia M., Gottsmann J., Carbone D., Fernándes J. 4D volcano gravimetry. *Geophysics*, vol.73, Nr. 6, (2008).
14. Van Camp M. Measuring seismic normal modes with the GWR C021 superconducting gravimeter. *Physics of the Earth and Planetary Interiors*, vol. 116, pp. 81–92, (1999).
15. Heki K., Matsuo K. Coseismic gravity changes of the 2010 earthquake in central Chile from satellite gravimetry. *Geophysical Research Letters*, vol. 37, L24306, (2010).
16. Dziewonski A. M., Anderson D. L. Preliminary reference Earth model. Physics of the Earth and Planetary Interiors, vol. 25, 297–356, (1981).
17. Stacey F. D. et al. Constraint on the planetary scale value of the Newtonian gravitational constant from the gravity profile within a mine. *Phys. Rev.* D23, 1683, (1981).
18. Holding S. C., Stacey F. D., Tuck G. J. Gravity in mines. An investigation of Newton's law. *Phys. Rev.*, D 33, 3487 (1986).
19. Stacey F. D. Gravity. *Science Progress*, vol. 69, No. 273, pp. 1–17, (1984).
20. Quinn T., Speake C. The Newtonian constant of gravitation – a constant too difficult to measure? An introduction. Phil. Trans. Royal Society A 372, 20140253.
21. Yanpeng Li. A Possible Exact Solution for the Newtonian Constant of Gravity. *British Journal of Mathematics and Computer Science*, vol. 19(3), 1–25, (2016).
22. Pound R. V., Rebka Jr. G. A. Gravitational Red-Shift in Nuclear Resonance. *Physical Review Letters*, 3 (9): 439–441, (1959).
23. Müller On the Acceleration of Free Fall inside Polyhedral Structures. *Progress in Physics*, 2018, vol. 14, 220–225.
24. Dermott S. F. On the origin of commensurabilities in the solar system - II: The orbital period relation. *Mon. Not. R. Astron. Soc.*, 141 (3): 363–376, (1968)
25. Lovis C. et al. The HARPS search for southern extra-solar planets. *Astronomy and Astrophysics*, manuscript Nr. HD10180, ESO, (2010)
26. Colombo G. Rotational Period of the Planet Mercury. *Letters to Nature*, Nr. 5010, p. 575, (1965)
27. Müller H. On the Cosmological Significance of Euler's Number. *Progress in Physics*, 2019, v. 15, 17–21.
28. Dombrowski K. Rational Numbers Distribution and Resonance. *Progress in Physics*, 2005, v. 1, no. 1, 65–67.
29. Müller H. The Physics of Transcendental Numbers. *Progress in Physics*, 2019, vol. 15, 148–155.
30. Müller H. Global Scaling. The Fundamentals of Interscalar Cosmology. *New Heritage Publishers*, Brooklyn, New York, USA, ISBN 978-0-9981894-0-6, (2018).
31. Müller H. Fractal Scaling Models of Natural Oscillations in Chain Systems and the Mass Distribution of Particles. *Progress in Physics*, 2010, v. 6, 61–66.
32. Müller H. Scale-Invariant Models of Natural Oscillations in Chain Systems and their Cosmological Significance. *Progress in Physics*, 2017, v. 13, 187–197.
33. Khintchine A.Ya. Continued fractions. University of Chicago Press, Chicago, (1964).
34. Hilbert D. Über die Transcendenz der Zahlen e und π. *Mathematische Annalen*, 43, 216–219, (1893).
35. Panchelyuga V.A., Panchelyuga M. S. Resonance and Fractals on the Real Numbers Set. *Progress in Physics*, 2012, v. 8, no. 4, 48–53.
36. Müller H. Quantum Gravity Aspects of Global Scaling and the Seismic Profile of the Earth. *Progress in Physics*, 2018, vol. 14, 41–45.
37. Müller H. Emergence of Particle Masses in Fractal Scaling Models of Matter. *Progress in Physics*, 2012, v. 8, 44–47.
38. Ries A. Bipolar Model of Oscillations in a Chain System for Elementary Particle Masses. *Progress in Physics*, 2012, vol. 4, 20–28.
39. Ries A. Qualitative Prediction of Isotope Abundances with the Bipolar Model of Oscillations in a Chain System. *Progress in Physics*, 2015, vol. 11, 183–186.
40. Ries A., Fook M. Fractal Structure of Nature's Preferred Masses: Application of the Model of Oscillations in a Chain System. *Progress in Physics*, 2010, vol. 4, 82–89.
41. Tanabashi M. et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018), www.pdg.lbl.gov
42. Astrophysical constants. Particle Data Group, pdg.lbl.gov
43. Planck M. Über Irreversible Strahlungsvorgänge. *Sitzungsbericht der Königlich Preußischen Akademie der Wissenschaften*, 1899, v.1, 479–480.
44. Müller H. Gravity as Attractor Effect of Stability Nodes in Chain Systems of Harmonic Quantum Oscillators. *Progress in Physics*, 2018, vol. 14, 19–23.
45. Müller H. Global Scaling of Planetary Atmospheres. *Progress in Physics*, 2018, v. 14, 66–70.
46. NASA Planetary Fact Sheet - Metric (2019).
47. Bills B. G., Ray R. D. Lunar Orbital Evolution: A Synthesis of Recent Results. *Geophysical Research Letters*, v. 26, Nr. 19, pp. 3045–3048, (1999).
48. Catalog of Exoplanets. Observatoire de Paris, http://exoplanet.eu/catalog/
49. Jacobson R. A. et al. The gravity field of the saturnian system from satellite observations and spacecraft tracking data. *The Astronomical Journal*, vol. 132, 2520–2526, (2006).
50. Jacobson R. A. The orbits of the uranian satellites and rings, the gravity field of the uranian system. *The Astronomical Journal*, vol. 148, 76–88, (2014).
51. Petropoulos A. E. Problem Description for the $6^t h$ Global Trajectory Optimisation Competition. *Jet Propulsion Laboratory*, California Institute of Technology, (2012).

The Role Played by Plasma Waves in Stabilizing Solar Nuclear Fusion

Tianxi Zhang

Department of Physics, Chemistry, and Mathematics, Alabama A & M University, Normal, Alabama 35762, USA.
E-mail: tianxi.zhang@aamu.edu

Since the wave function of two-scattering protons has been used for that of diproton or helium-2 in the conventional analysis with Fermi theory, the probability for a diproton to form a deuteron via a β^+-decay has been extremely under calculated. This implies that the rareness of β^+-decay in diprotons is not rare enough to inhibit the solar nuclear fusion. To meet the observed rate of solar nuclear fusion, the core of the Sun must involve another significant physical effect to inhibit solar nuclear fusion. This study finds that plasma waves can play this role, because they significantly reduce the electric permittivity of the core plasma and thus extremely raise the Coulomb barrier or shift the Gamow peak to a higher energy of particles. It is shown that, if the frequency of plasma waves that are globally generated in the core plasma of turbulences is about 1.28 times the plasma frequency, the Sun can have the actual fusion rate or shine on at the currently observed luminosity. Therefore, in addition to the quantum tunneling effect and rareness of β^+-decay, plasma waves can also play an essential role in the solar nuclear fusion and power emission. The result of this study may also give implications to supernova explosion, missing solar neutrino, and plasma nuclear fusion in laboratory.

1 Introduction

The Sun is a giant natural fusion reactor [1]. It smashes about 3.6×10^{38} hydrogen nuclei or protons per second to produce helium nuclei or α-particles, while releasing nuclear power of 3.85×10^{26} W. This nuclear fusion process occurring in the core of the Sun has been comprehensively investigated for many decades based on the well-developed stellar nucleosynthesis and quantum physics. It is well known that, in the dense and hot core of the Sun with ~1.5 keV (or $\sim 1.67 \times 10^7$ K) temperature and Boltzmann-Maxwell's distribution of the core's total 1.2×10^{56} protons, there should not be any proton able to overcome the 820 keV (or 9.5×10^9 K) Coulomb barrier to make the fusion reactions occur.

According to Gamow's quantum tunneling probability [2] however, the energy region where nuclear reactions are most like ly to occur (i.e. the Gamow peak) is around 10^8 K. This allows one part per million of the core's total 1.2×10^{56} protons to penetrate the Coulomb barrier. With this probability of barrier tunneling or penetration, the high ion-collision frequency of 20 terahertz means that the core of the Sun fuses all its protons within the order of only microseconds (i.e. a rate of 10^{63} s^{-1}, 25 orders of magnitude higher than the actual reaction rate) and thus would instantaneously explode. It is generally believed that the major reasons why the Sun does not instantaneously blow up are (1) the difficulty of double proton (also called diproton) formation (estimated to be lowered only by $\sim 10^{-6}$ according to the Gamow tunneling probability), (2) the rareness of β^+-decay from diprotons (needed to be lowered by $\sim 10^{-25}$ according to the Sun's actual luminosity) and (3) the squeezing of the Sun's strong gravity.

However, in the conventional analysis and calculation of the Fermi theory of the β^+-decay, the significant wave function of two-scattering protons was usually used for the inefficient wave function of the diproton outside the potential energy well [3]. This is not physical and greatly weakens the wave function of the diproton inside the potential energy well, so that leads to the probability for a diproton to form a deuteron via a β^+-decay to be extremely under calculated [4]. In other words, the rareness of β^+-decay in diprotons may not be rare enough to inhibit the solar nuclear fusion or lower the fusion rate by 25 orders of magnitude, in order to stop the Sun's instantaneous explosion and have the currently observed luminosity. The quantum tunneling effect allows many diprotons formed in the Sun's core, but the probability for a diproton to form a deuteron via a β^+-decay may not be lower than that for a diproton to separate back to two protons by 25 orders of magnitude. Observations have only given an upper bound that a diproton (or helium-2 nucleus) gets β^+-decay by less than one per ten thousands, i.e. < 0.01% [5].

In this paper, we propose a new mechanism of inhibition that can significantly reduce the fusion reaction rate and thus effectively prevent the Sun from an instantaneous explosion. We suggest that the core of the Sun involves a significant physical effect or inhibitor called plasma oscillation or wave, which significantly reduces the electric permittivity of the core plasma. A significantly reduced electric permittivity will greatly raise the Coulomb barrier as well as efficiently lower the Gamow tunneling probability. These changes lead to greatly shift the Gamow peak to the region of higher energies of particles. Quantitative study in this paper indicates that if the frequency ω of plasma oscillations or waves that are globally generated in the core plasma of turbulences is about 1.28 times the plasma frequency ω_p, the Sun can have the actual fusion rate or shine on at the currently observed luminosity.

Therefore, in addition to the quantum tunneling effect, the plasma oscillations may play also an essential role in the Sun's nuclear fusion and power emission. The quantum tunneling effect makes the fusion to occur, while the plasma oscillations in association with the weak β^+-decay of diprotons guarantees that the Sun will not explode. We also suggest that a supernova explosion occurs when plasma oscillations in the core of a star at the end of its life are significantly weakened in intensity or changed in frequency, that causes the heavy ion fusion to be significantly speeded up and the huge amount of energies and neutrinos to be instantaneously emitted. The result of this study also gives important implications to plasma nuclear fusion in laboratory and the solar neutrino missing problem.

2 Coulomb barrier and solar nuclear power emission

The measurement of power emission indicates that the luminosity of the Sun at present is about 3.85×10^{26} W, which can be calculated from

$$L_\odot = 4\pi R^2 \sigma T^4, \tag{1}$$

where $R = 7 \times 10^8$ m is the radius of the Sun, $\sigma = 5.67 \times 10^{-8}$ W/(m^{2}K^4) is the Stefan-Boltzmann constant, and $T = 5778$ K is the surface temperature of the Sun. At this luminosity, the Sun's gravitational energy, determined by

$$U = \frac{3GM^2}{5R} \simeq 2.3 \times 10^{41}\ \mathrm{J}, \tag{2}$$

can only let it shine about $U/L_\odot \sim 19$ million years, which is the thermal or Kelvin-Helmholtz timescale determined by $K/L_\odot$ and is much shorter than the actual Sun's lifetime. Here $G = 6.67 \times 10^{-11}$ N m^2/kg^2 is the gravitational constant, $M = 1.99 \times 10^{30}$ kg is the mass of the Sun, K is the internal energy or the total kinetic energy of particles in the Sun, determined by

$$K = \frac{3}{2}\, k_B N T_{\mathrm{core}} \simeq 4.1 \times 10^{41}\ \mathrm{J}, \tag{3}$$

with $k_B = 1.38 \times 10^{-23}$ the Boltzmann constant, $N = M/m_p$ the total number of protons within the Sun, $m_p = 1.67 \times 10^{-27}$ kg the mass of the proton, and $T_{\mathrm{core}} = 1.67 \times 10^7$ K the temperature of the core of the Sun. It should be noted here that the hot core of the Sun is about 1/3 of its diameter or 1/10 of its mass, which means that the internal energy of the Sun should be several times less than that given by (3).

The total number of protons in the core of the Sun is given by

$$N_0 = \frac{1}{10}\,\frac{M}{m_p} \sim 1.2 \times 10^{56}, \tag{4}$$

or number density to be $n_0 \sim 2.2 \times 10^{30}$ m^{-3}. It is the number or number density of protons available for fusion and the Sun should be mainly powered by nuclear fusion. According to nuclear physics, every time four protons are fused to form one helium, the reactions produce two neutrinos, two positrons, and two photons, and release in total a net energy of $E_{4p} \sim 27$ MeV from the deficit of ~3% masses of four protons. The energy from the fusion of all protons in the core of the Sun, calculated by

$$E = \frac{1}{10}\,\frac{M}{4m_p}\,E_{4p} \simeq 1.3 \times 10^{44}\ \mathrm{J}, \tag{5}$$

can run the Sun at the present rate of emission for about 10 billion years. On the other hand, to have the present energy emitting rate, the Sun needs to fuse its protons at a rate of about

$$\frac{dN_0}{dt} = \frac{4L_{Sun}}{E_{4p}} \simeq 3.6 \times 10^{38}\ \mathrm{s}^{-1} \tag{6}$$

protons in one second.

In order to fuse protons, the extremely high Coulomb barrier between them, determined by

$$U_C = \frac{q_p^2}{4\pi\epsilon_0 d_p} \simeq 8.2 \times 10^2\ \mathrm{keV}\ \text{ or }\ 9.5 \times 10^9\ \mathrm{K}, \tag{7}$$

must be overcome [6]. Here $q_p = 1.6 \times 10^{-19}$ C is the proton's electric charge (equal to the fundamental unit of charge e), $\epsilon_0 = 8.85 \times 10^{-12}$ C^2/(m^{2}N) is the electric permittivity constant in free space, and $d_p = 1.76 \times 10^{-15}$ m is the diameter of a proton. Since the average kinetic energy of protons in the Sun's core with temperature 1.67×10^7 K, $K = (3/2) k_B T_{core} = 2.16$ keV, is about 383 times lower than the Coulomb barrier between protons, and it must be very hard to have protons to be able to climb over the Coulomb barrier. According to the Boltzmann-Maxwellian distribution function [7, 8], we have the number of protons with velocity in the range v - v+dv to be given by

$$dN = N_0 \left(\frac{m}{2\pi k_B T}\right)^{3/2} 4\pi v^2 \exp\left(-\frac{mv^2}{2k_B T}\right) dv, \tag{8}$$

or with energy in the range of E - E+dE to be given by,

$$dN = N_0 \frac{2\pi}{(\pi k_B T)^{3/2}} \sqrt{E} \exp\left(-\frac{E}{k_B T}\right) dE. \tag{9}$$

Here N_0 is the total number of all particles. Then, the number of protons with energy above the Coulomb barrier U_C can be found by integrating the function (9) with respect to the energy (E) in the range from U_C to infinity as

$$\begin{aligned} N_C &= N_0 \int_{U_C}^{\infty} \frac{2\pi}{(\pi k_B T)^{3/2}} \sqrt{E} \exp\left(-\frac{E}{k_B T}\right) dE \\ &= \frac{2N_0}{\sqrt{\pi}} \sqrt{\frac{U_c}{k_B T}} \exp\left(-\frac{U_c}{k_B T}\right) + N_0 \operatorname{erfc}\left(\sqrt{\frac{U_c}{k_B T}}\right). \end{aligned} \tag{10}$$

Considering the ion collision frequency in the hot core of the Sun to be calculated by

$$\nu_i = 4.8 \times 10^{-8}\, Z_i^4 \mu^{-1/2} n_i \ln\Lambda\, T_i^{-3/2}\ \mathrm{s}^{-1}, \tag{11}$$

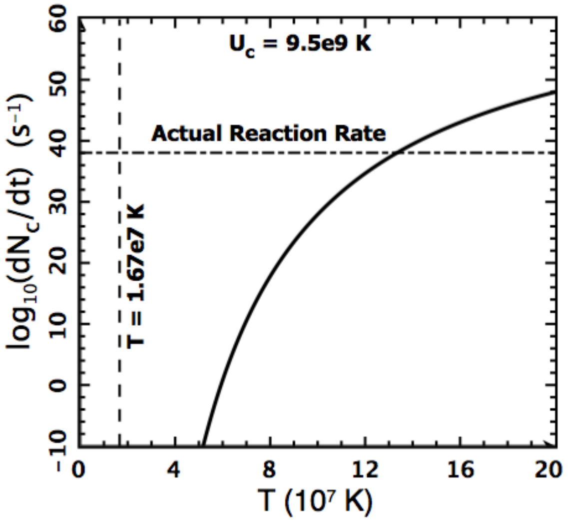


Fig. 1: The reaction rate of protons is plotted as a function of the Sun's core temperature in the case of without considering the quantum tunneling effect. The result indicates that no nuclear fusion can actually occur.

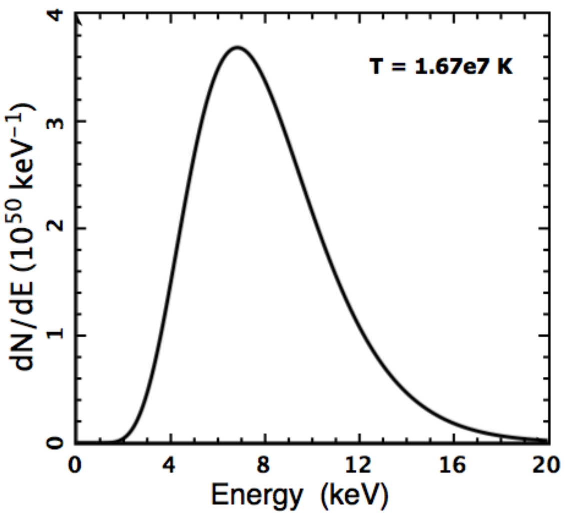


Fig. 2: Energy spectrum of protons that can penetrate the Coulomb barrier for fusion. The number of tunneling protons per unit energy in the core of the Sun is plotted as a function of the energy. The maximum is usually called the Gamow peak, which is located near the energy of about 7 keV.

where Z_i is the ion charge state, μ is the ion-proton mass ratio, n_i is the number of ions per cubic centimeter, $\ln\Lambda$ is the Coulomb logarithm with a convenient choice to be 10, and T_i is the ion temperature in units of eV. For protons in the Sun's core, the collision frequency can be $\nu_p \sim 2\times10^{13}$ Hz. The reaction rate of protons that can climb over the Coulomb barrier can then be estimated by

$$\frac{dN_C}{dt} = N_c\,\nu_p\,\mathrm{s}^{-1}. \tag{12}$$

Fig. 1 plots this reaction rate of protons as a function of the core temperature. It is seen that the reaction rate of the protons is about zero (many orders of magnitude less than 10^{-10} s^{-1}), so that no nuclear fusion occurs in the core of the Sun if the core temperature is equal to the conventional value $T_{core} = 1.67\times10^7$ K. For the reaction rate of protons to be the actually observed rate of 3.6×10^{38} protons per second, the Sun's core temperature must be about 1.3×10^8 K or above. Therefore from classical physics, solar nuclear fusion will hardly occur.

3 Quantum tunneling effect on solar nuclear reaction

Quantum tunneling effect plays an essential role in solar nuclear fusion. According to the Gamow tunneling probability [2], given by

$$P_g = \exp\left(-\sqrt{\frac{E_g}{E}}\right), \tag{13}$$

one can determine the number of protons with energy between E and $E+dE$ that can tunnel through or penetrate the Coulomb barrier as

$$\begin{aligned} dN_g &= P_g dN \\ &= N_0\,\frac{2\pi}{(\pi k_B T)^{3/2}}\,\sqrt{E}\exp\left(-\frac{E}{k_B T}-\sqrt{\frac{E_g}{E}}\right)dE, \end{aligned} \tag{14}$$

where E_g is the Gamow energy determined by

$$E_g = 2m_r c^2(\pi\alpha Z_a Z_b)^2. \tag{15}$$

Here m_r is the reduced mass of the nuclei, c is the speed of light, Z_a and Z_b are the ionization states of the nuclei, and $\alpha = e^2/(2\epsilon_0 hc)$ is the fine-structure constant.

The distribution (14) for the number of tunneling protons with respect to the energy exhibits a maximum called the Gamow peak that has energy to be significant (about 120 times) less than the Coulomb barrier, so that the quantum tunneling effect greatly enhances the reaction rate in the core of the Sun. To see in more details the increase of the tunneling probability, we plot in Fig. 2 the Gamow peak for the Sun's core with temperature 1.67×10^7 K. The energy of the peak is around 7 keV and the height of the peak is around 3.7×10^{50} protons per keV.

Both the height and energy of the Gamow peak depend on the temperature of the Sun's core. Evaluating the extreme value of (14), we can obtain the energy of Gamow peak as a function of the core's temperature and other parameters or constants as the following implicit equation

$$1-\frac{2E_p}{k_B T}+\sqrt{\frac{E_g}{E_p}} = 0. \tag{16}$$

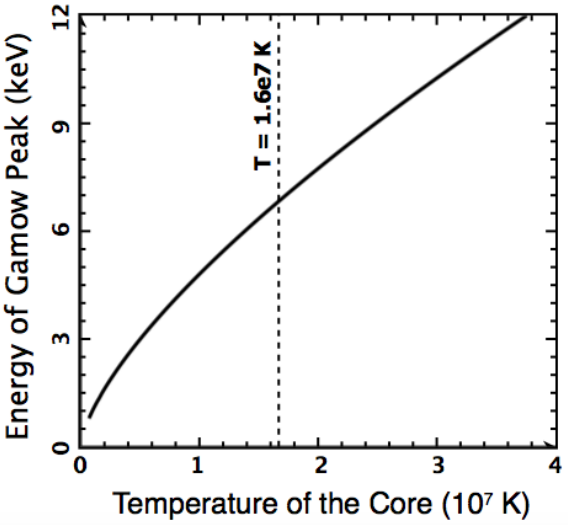


Fig. 3: The energy of the Gamow peak is plotted as a function of the temperature of the core.

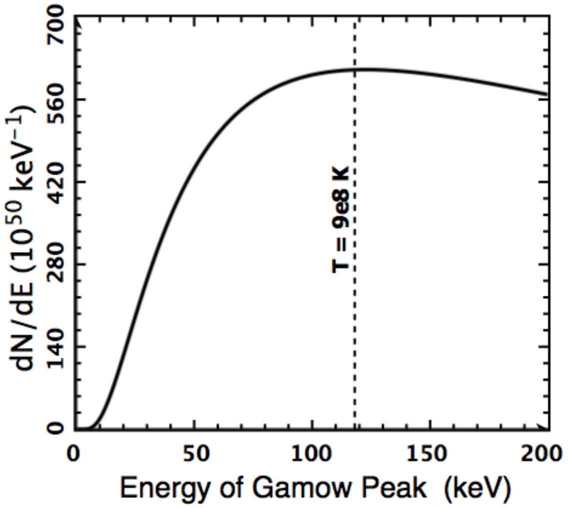


Fig. 4: The number of protons per unit energy is plotted as a function of the energy of the Gamow peak, which increases with the temperature core.

Substituting the energy of the Gamow peak (E_p) back into (14), we can determine the height of the Gamow peak as a function of the core's temperature and other parameters or constants. Fig. 3 plots the energy of the Gamow peak as a function of the temperature of the core. It is seen that the energy of the Gamow peak increases as the temperature of the core increases. The Gamow peak is at about 7 keV if the core temperature is 1.67×10^7 K and increases to 10 keV when the core temperature increases to 3×10^7 K. Fig. 4 plots the number of tunneling protons per unit energy (i.e. per keV) as a function of the energy of the Gamow peak, which increases as the temperature of the core increases as shown in Fig. 3. This further shows that the number of tunneling protons per unit energy reaches a maximum ($\sim 6 \times 10^{52}$ keV^{-1}) when the energy of the Gamow peak is about 120 keV (or the temperature of the core is about 0.9 billion Kelvins). In the Sun's core temperature of 1.67×10^7 K, the energy of the Gamow peak is only 7 keV and the maximum number of tunneling protons is about 3.6×10^{50} keV^{-1}. Based on this peak of the maximum number of tunneling nuclei, we can find the maximum reaction rate as a function of the energy of the Gamow peak or the temperature of the core. This result may be important to optimize plasma fusion in the laboratory.

Then, the number of protons that can penetrate or tunnel through the Coulomb barrier can be found by integrating the function (14) with respect to the energy (E) in the range from zero to infinity as

$$\begin{aligned} N_g &= \int_0^\infty P_g dN \\ &= \lim_{E_2 \to \infty} \int_0^{E_2} \frac{2\pi\sqrt{E}}{(\pi k_B T)^{3/2}} \exp\left(-\frac{E}{k_B T} - \sqrt{\frac{E_g}{E}}\right) dE. \end{aligned} \tag{17}$$

Multiplying N_g with the collision frequency, we obtain the reaction rate of nuclear fusion with the quantum tunneling effect as

$$\frac{dN_g}{dt} = N_g \nu_p. \tag{18}$$

To see the reaction rate quantitatively, we plot in Fig. 5 the reaction rate (18) as a function of the upper energy of the integration (E_2), which should approach infinity (or a value that is big enough, e.g. 30 keV). For the core of the Sun, the reaction rate saturates at $\sim 2 \times 10^{63}$ protons per second when the upper energy of the integration is $E_2 \gtrsim 30$ keV. This reaction rate is an order of magnitude 25 times higher than the actual reaction rate. Without a significant inhibitor to greatly slow down the reactions, the Sun should have instantaneously exploded.

4 Plasma oscillation effect on solar nuclear fusion

Plasma oscillations or waves can be considered as a great inhibitor for the solar nuclear reaction, because the dielectric constant of plasma with plasma oscillations or waves is given by [9]

$$\epsilon_r = 1 - \frac{\omega_p^2}{\omega^2}, \tag{19}$$

where ω_p is the plasma frequency defined by

$$\omega_p = \sqrt{\frac{n_e e^2}{\epsilon_0 m_e}}, \tag{20}$$

and ω is the frequency of plasma waves generated in the core by the oscillations of free electrons. Eq. (19) indicates that

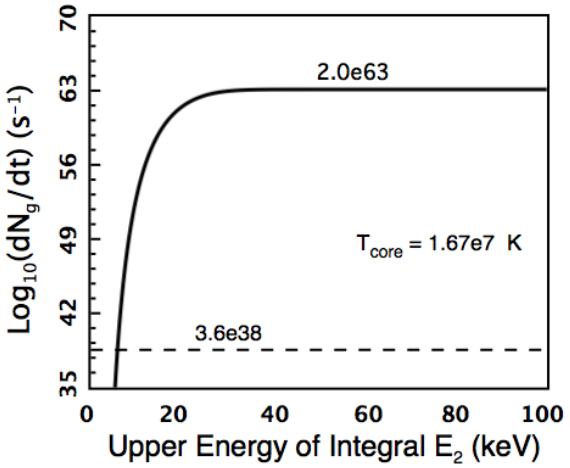


Fig. 5: The reaction rate of protons in the core of the Sun. The number of adequate collisions per second between protons is plotted as a function of the upper energy of the integration.

plasma oscillations or waves can make the dielectric constant to be less than unity and hence raises the Coulomb barrier and increases Gamow energy or reduces quantum tunneling probability. Increases of both the Coulomb barrier and the Gamow energy can greatly reduce the fusion reaction rate.

There are several types of plasma waves that can be initiated by electron oscillations [10] such as electrostatic Langmuir waves [11] with the dispersion relation given by

$$\omega^2 = \omega_p^2 + 3k^2 v_{Te}^2, \tag{21}$$

where v_{Te} is the thermal velocity of electrons and k is the wavenumber. Fig. 6 shows the dispersion relation of the Langmuir waves by plotting the wave frequency as a function of the wavenumber. It is seen that the frequency is about 1.28 times the plasma frequency when the wavenumber is about $k \sim 10^{9.4} \sim 2.5 \times 10^9$ m^{-1}. It is about half of the wavenumber of blackbody radiation at the peak, $k_{bb} = T_{core}/(2.9 \times 10^{-3}) \sim 5.8 \times 10^9$ m^{-1} $\sim 2.3k$ and also half of the Debye wavenumber, $k_d = [n_e e^2/(\epsilon_0 k_B T_{core})]^{1/2} \sim 5.0 \times 10^9$ m^{-3} $\sim 2k$. In the core of the Sun, we have $\omega_p \sim 3.6 \times 10^{18}$ Hz, i.e. in the X-ray frequency range.

To see the plasma oscillation effect on the solar nuclear fusion, we plot in Fig. 7 the reaction rate (18) as a function of the upper energy of the integration in three cases with the frequencies of plasma oscillations or waves given by $\omega/\omega_p = 1.25, 1.28, 1.32$, respectively. For the core of the Sun with $\omega/\omega_p \sim 1.28$, the reaction rate saturates at $\sim 3.6 \times 10^{38}$ protons per second when the upper energy of the integration is $E_2 \gtrsim 80$ keV. This reaction rate is in magnitude about the order of the actual reaction rate. Slightly varying the frequency, we have a reaction rate quite different. In general, the higher the frequency of the plasma waves, the weaker the effect of plasma oscillations on the nuclear fusion is. It is an extremely efficient inhibitor of the solar nuclear fusion.

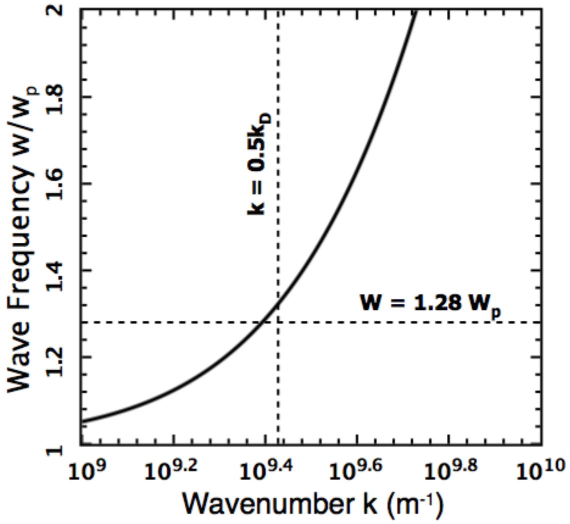


Fig. 6: The dispersion relation of the plasma Langmuir waves. The frequency of the waves is plotted as a function of the wavenumber. It is seen when the wavenumber is about half of the Debye wavenumber or wavelength is about the diameter or twice the radius of the Debye sphere.

5 Discussions and conclusions

At the end of its life, a star runs out proton-proton fusion and thus varies the plasma oscillations, which causes this efficient inhibitor to be ineffective. With this reason, we suggest that supernova explosions may occur when plasma oscillations in the core of a star at the end of its life are significantly weakened in intensity or changed in frequency that cause the heavy ion fusion to be significantly speeded up and out of control and the huge amount of energies and neutrinos to be instantaneously emitted. This study of the role of plasma oscillation played in solar nuclear fusion provides us an alternative mechanism for supernova explosions, in addition to the previously proposed and developed models of supernova explosions driven by magnetohydrodynamtic (MHD) rotation [12], acoustic waves [13], neutrinos [14], and gravitational field shielding [15].

The plasma oscillations or waves with frequency about 1.28 times the plasma frequency can reduce the electric permittivity or the dielectric constant by a factor of one third in comparison with free space. The effective refractive index of plasma is given by $n = \sqrt{\epsilon_r} \sim 0.6$. Postulating the mass energy conversion in plasma by $E = mc^2/n^2$ leads to the deficit of 3% proton masses in fusion that can produce three times the nuclear energy. Then, having the same luminosity, the Sun only needs to fuse one third of the early suggested num-

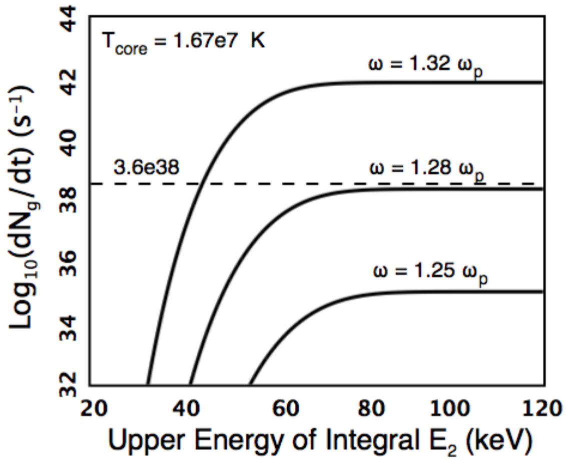


Fig. 7: The reaction rate of protons in the core of the Sun. The number of adequate collisions per second between protons is plotted as a function of the upper energy of the integration. Here the plasma oscillation effect on the reaction rate is included.

ber of protons, i.e. $\sim 1.2 \times 10^{38}$ protons per second. This result provides us an alternative of quantitatively explaining the missing two thirds of the solar neutrinos [16]. The Sun's lifetime is thus tripled, to be over thirty billion years.

Plasma oscillations with appropriate frequency of disturbances may also affect the nuclear reactions of plasma fusion in the laboratory. Above the plasma frequency ($\omega > \omega_p$), plasma oscillations would reduce the reaction rates and hence make the fusion hard to occur. Below the plasma frequency ($\omega < \omega_p$), however, plasma oscillations can lead to a negative dielectric constant, which switches the Coulomb interaction between nuclei to be an attractive force from a repulsive one. In this case, the Coulomb barrier disappears and nuclei fuse freely. Therefore, the result of this study also gives an important implication to plasma nuclear fusion in the laboratory. Regarding plasma fusion in the laboratory, the author has recently developed two new mechanisms: (1) plasma fusion at some keV with extremely heated ^{3}He ions or tritons [17–19]; (2) plasma fusion with Coulomb barrier lowered by scalar field [20].

As a consequence of this study, except for the conventional inhibitor of unlikely β^+-decay from diprotons, we find that plasma oscillations or waves can be an efficient inhibitor for the solar nuclear fusion, as it significantly reduces the electric permittivity of the core plasma and thus extremely raises the Coulomb barrier or shifts the Gamow peak to a higher energy of particles. When the frequency of plasma oscillations or Langmuir waves that are globally generated in the core plasma of turbulences is about 1.28 times the plasma frequency, the Sun can have the actual fusion rate or shine on at the currently observed luminosity. Therefore, in addition to the quantum tunneling effect, the weak β^+-decay, the plasma oscillations play also an essential role in the Sun's nuclear fusion and power emission.

Acknowledgements

The author acknowledges Bing Li for support and also thanks reviewers and editors for improving the manuscript quality.

Received on February 2, 2020

References

1. Zirin H. Astrophysics of the Sun. Cambridge Univ. Press, Cambridge, 1988.
2. Gamow G. Zur Quantentheorie des Atomkernes. *Z. Physik*, 1928, v. 51, 204–212.
3. Bethe H. T. The Formation of Deuterons by Proton Combination. *Physical Review*, 1938, v. 54, 248-254
4. Zhang T. X. A Critical Issue on Solar Nuclear Fusion. 2021, to be submitted
5. en.wikipedia.org/wiki/Isotopes_of_helium.
6. Coulomb P. M. Histoire de l'Académie Royale des Sciences. 1875, 569–577.
7. Maxwell J. C. On the Motions and Collisions of Perfectly Elastic Spheres. *Philosophical Magazine and Journal of Science*, 1860, v. 19, 19–32.
8. Boltzmann L. Weitere Studien uber das Warmegleichgewicht unter Gasmolekulen. *Sitzungsber. Kais. Akad. Wiss.*, 1872, v. 66, 275–370.
9. Fitzpatrick R. Classical Electromagnetism. http://farside.ph.utexas.edu /teaching, 2006.
10. Stix T. H. Waves in Plasmas. Springer, AIP Press, 1992.
11. Langmuir I. The Interaction of Electron and Positive Ion Space Charges in Cathode Sheaths. *Physical Review*, 1929, v. 33, 954
12. Meier D. L., Epstein R. I., Arnett D., Schramm D. N. Magnetohydrodynamic Phenomena in Collapsing Stellar Cores. 1976, *Astrophysical Journal*, 1976, v. 204, 869–878.
13. Burrows A., Livne E., Dessart L., Ott C. D., Murphy J. Features of the Acoustic Mechanism of Core-Collapse Supernova Explosions. *Astrophysical Journal*, 2007, v. 655, 416–433.
14. Colgate S. A., White R. H. The Hydrodynamic Behavior of Supernovae. *Astrophysical Journal*, 1966, v. 143, 626–681.
15. Zhang T. X. Gravitational Field Shielding and Supernova Explosions. *Astrophysical Journal Letters*, 2010, v. 725, L117–L121.
16. Davis R. Jr., Harmer D. S., Hoffman K. C. Search for Neutrinos from the Sun. *Physical Review Letters*, 1968, v. 20, 1205–1209.
17. Zhang T. X. Two-Stage Heating Mechanism for Plasma Fusion at 10 MK. *Proceedings of IEEE 25th Symposium on Fusion Engineering (SOFE*, 2013, 978-1-4799-0171-5/13.
18. Zhang T. X.,Ye M. Y. Plasma fusion at 10 MK with Extremely Heated ^{3}He Ions. *IEEE Transactions on Plasma Science*, 2014, v. 42, 1430–1437.
19. Zhang T. X. Plasma fusion of Deuterons at Kiloelectron Volts with Extremely Heated Tritons. *IEEE Transactions of Plasma Science*, 2020, v. 48, 2884–2891.
20. Zhang T. X. Nuclear Fusion with Coulomb Barrier Lowered by Scalar Field. *Progress in Physics*, 2019, v. 15, 191–196.

Antarctic Circumpolar Current: Driven by Gravitational Forces?

Franklin Potter

Sciencegems, 8642 Marvale Drive, Huntington Beach, CA 92646 USA.
E-mail: frank11hb@yahoo.com
ORCID: 0000-0002-9440-2291

The Antarctic Circumpolar Current (ACC) is the largest ocean current, one that travels west to east at a velocity of about 2 m/s greater than the Earth's rotation velocity at latitudes from 40ºS to about 60ºS. Simple models of the winds driving this current consider a linear relationship between the wind strength and the water transport. However, the behavior is much more complex. The ultimate energy source driving the winds and this current remains to be identified. I investigate whether a gravitational force dictated by Quantum Celestial Dynamics (QCM) is the true energy source that maintains the ACC.

1 Introduction

The Antarctic Circumpolar Current (ACC), the largest ocean current on Earth, flows west to east at about 2 m/s faster than the Earth's rotation at its latitude of about 40ºS to about 60ºS near the Antarctic continent [1, 2], as shown in Figure 1. Its mean transport is estimated to be about 134 sverdrup, i.e., 134×10^6 m^3/s. There are two different atmospheric winds to consider: the winds along the ACC and the winds along the contours of Antarctica, with variations in both able to cause robust changes in ACC transport. They are considered to be the major driving force of this enormous water current.

But the ACC current extends to the ocean floor, with a strong current velocity of about 2 cm/s at a depth of 3000 meters [3, 4]. So this approach becomes quite complicated by involving thermodynamic mixing vertically and horizontally, various wind strength and direction changes, Coriolis force effects, eddies, etc.

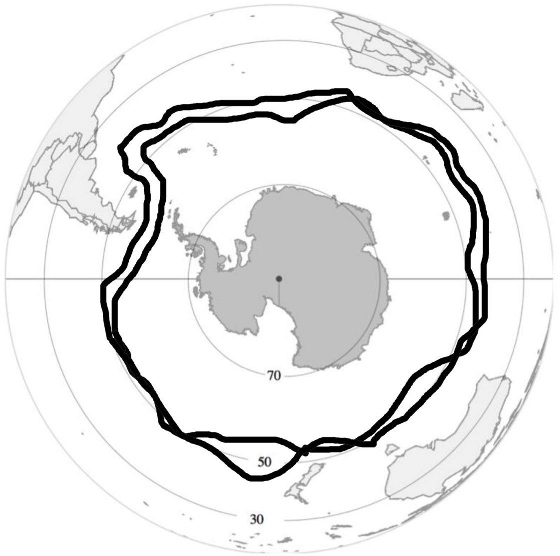


Fig. 1: Antarctic Circumpolar Current moving west to east faster than the Earth's rotation, showing its deviations from a circular path with many latitude variations.

Ultimately, one might expect to identify a powerful and consistent energy source that would be capable of forcing such a large water transport at all depths as well as help drive the winds in the atmosphere.

Herein I apply Quantum Celestial Mechanics (QCM) to the binary system of the rotating Earth and the orbiting Moon, both objects providing the total system vector angular momentum required by QCM [5] to determine its gravitational stationary energy states exhibiting the quantization of angular momentum per unit mass. I can use the familiar general relativistic Schwarzschild metric because the QCM equilibrium radii r_{eq} are much larger than the 9 millimeter Schwarzschild radius r_g of the Earth. These QCM states at specific equilibrium radii in the plane of the Equator are assumed to define rotational cylinders co-axial to the Earth's rotation axis that intersect the Earth's surface. In particular, I am interested in determing whether the QCM angular momentum quantization per unit mass approach can be the source of the driving force responsible for the Antarctic Circumpolar Current.

2 QCM brief history review

In 2003 Howard G. Preston and I introduced [5] Quantum Celestial Mechanics (QCM) to explain the spacings of planetary orbits in the Solar System and in all known exoplanetary systems. In the Schwarzschild metric, the quantization of the total angular momentum per unit mass in a gravitationally bound system constrains the possible orbital radii to specific allowed values determined by quantization integers.

At that time, we were not successful in finding a system that could be a definitive test of QCM. Unfortunately, there existed no gravitationally bound system with three or more celestial objects for which the angular momentum was known to within 10%, not for the Solar System nor for the Jovian planets and their satellites. Therefore, we proposed several laboratory experiments to test for a repulsive gravitational QCM force, including the response of two pendulums in a microwave vacuum chamber and of the response of one LIGO interferometer to the slow one rotation per hour spin of a 10 kg mass several meters distant. Neither tests were approved.

However, the 2015 New Horizons flyby of Pluto and its 5 moons did provide the data [6] for the definitive test of QCM, with the predicted QCM orbital constraint relation verified to within 2.4%.

The QCM gravitational wave equation derived from the general relativistic Hamilton-Jacobi equation is

$$g^{\alpha\beta}\frac{\partial^2\Psi}{\partial x^\alpha \partial x^\beta}+\frac{\Psi}{H}=0 \tag{1}$$

in which the scalar $\Psi = \exp[iS'/H]$, for S'= $S/\mu_i c$, with S the classical action, μ_i the mass of the particle acted upon, and c the speed of light in vacuum. The system scaling length is defined as H = $L_T/M_T c$, with L_T the total vector angular momentum for the system of total mass M_T.

This QCM gravitational wave equation is not quantum gravity. However, there is a relationship to the Schrödinger equation in quantum mechanics that was derived from the normal Hamilton-Jacobi equation using the transformation $\Psi = \exp[iS/h]$, with the classical action S and the universal Planck's constant h. Our H is not a universal constant.

The inherent generality and power of this gravitational wave equation arises from its dependence upon only two important physical parameters that characterize the gravitationally bound system: the total mass M_T and the total vector angular momentum L_T, both quantities defining H. In planetary systems, for example, the larger the value of H, the larger the spacings will be between the allowed QCM orbital equilibrium radii.

Successful applications of QCM have included the prediction of a Solar System total angular momentum of 1.86×10^{45} kg-m^2/s, most of which is contributed by the Oort Cloud at about 40,000 AU, a value about 50 times the listed angular momentum of the Sun plus the 8 planets [7]. Compared to all the known exoplanetary systems, our Solar System is unique because no other system exhibits such large planetary spacings that require this large total system angular momentum value.

Successful applications to galaxies and clusters of galaxies describe how QCM can fit their almost constant rotational velocities without invoking dark matter. Also, QCM was shown to be able to derive the MOND relation, which fits the galaxy rotational data extremely well and is considered a viable competitor to dark matter approaches [8].

A new interpretation [9] of the redshifts of light from distance sources in the Universe was introduced by applying the interior metric in a static Universe, thereby revealing a possible negative QCM gravitational potential that becomes more negative non-linearly from the observer, meaning that the light source is in a deeper negative gravitational potential for all observers. As such, the clocks at the light source tick slower than at the observer and the observed redshifts are purely gravitational redshifts. No dark energy is required to agree with the measured SNe 1a redshifts that have been interpreted as a recently accelerating Universe, and the Hubble value becomes distance dependent.

3 QCM Schwarzschild metric radial equation

Applying the general relativistic Schwarzschild metric to the QCM wave equation for radius values beyond r_g, dropping very small terms, and then evaluating the angular equations in spherical polar coordinates, leaves the radial r equation [5]

$$\frac{d^2\Psi}{dr^2}+\frac{2}{r}\frac{d\Psi}{dr}+\left[\frac{E}{\mu}+\frac{r_g c^2}{2r}-\frac{\ell(\ell+1)H^2c^2}{2r^2}\right]\Psi \simeq 0, \tag{2}$$

with ℓ the angular momentum integer from the θ and ϕ coordinates.

From the energy expression in the square bracket, the effective potential

$$V_{eff}=-\frac{r_g c^2}{2r}+\frac{\ell(\ell+1)H^2c^2}{2r^2}, \tag{3}$$

and the equilibrium radius for the QCM state ℓ is

$$r_{eq}=\ell(\ell+1)\frac{2H^2}{r_g}. \tag{4}$$

If one decides to use the Schwarzschild metric in cylindrical coordinates instead, then the product $\ell(\ell+1)$ usually becomes replaced by m^2, with m the integer for the ϕ direction quantization.

I will take this r_{eq} to be at the plane of the Equator for defining a cylinder co-axial with the Earth's rotation axis that extends in both directions to intersect the Earth's surface in North and South latitudes. Thus, by knowing the H and r_g values to calculate r_{eq}, one can predict the equilibrium radii of all the QCM states.

4 Results

4.1 Earth spin only

The total vector angular momentum of the Earth-Moon system is required by QCM. However, a preliminary simple calculation that considers just the rotation of the Earth about its axis is instructive.

The pertinent physical parameters of the Earth-Moon system are listed in Table 1, including the Earth's moment of inertia factor $\alpha = 0.827$ and the average angle factor β = Cos(23.4°) between the Earth's equatorial plane and the plane of the Moon's orbit. If only the Earth's spin angular momentum is considered, H = 3.26 meters, so

$$r_{eq}=\ell(\ell+1)\;2.36\times 10^3\ \mathtt{m}. \tag{5}$$

Beginning with the $\ell = 1$ state, all the r_{eq} values will be too small for any important relationship to the ACC around Antarctica.

Table 1: Earth–Moon QCM parameters.

Parameter	**Spin only**	**Earth–Moon**
Mass (10^{24} kg)	5.972	6.045
Radius (10^6 m)	6.37	385.0
Period (10^6 s)	0.08614	2.3605
α	0.827	—
β	—	0.918
L_T (10^{33} kg-m^2/s)	5.847	32.5
H (m)	3.26	17.94

4.2 Earth–Moon total angular momentun

The QCM wave equation requires the total vector angular momentum of the gravitationally bound system in its applications. The orbital vector angular momentum value for the Moon is the much larger contributor in the Earth–Moon system but varies considerably, repeating every 18.6 years, because the angle between the Moon's orbital plane and the Earth's equatorial plane reaches a maximum difference angle of 28º36' and a minimum of 18º20'.

Without accounting for this variation in the difference angle, the Moon's orbital motion would contribute about 2.91×10^{34} kg-m^2/s. Assuming an average difference angle of about 23.4º with repect to the Earth's equatorial plane, with cosine 23.4º = 0.918, the Moon's average vector orbital angular momentum contribution is about 2.67×10^{34} kg-m^2/s.

Therefore, the Earth-Moon H = 17.94 meters and

$$r_{eq} = \ell(\ell + 1)\ \ 71.52 \times 10^3\ \mathtt{m}. \tag{6}$$

The r_{eq} calculated values and their surface intersection latitudes for $\ell = 1$ to 9 are listed in Table 2.

The two QCM equilibrium radii r_{eq} for $\ell = 6$ and $\ell = 7$ intersect the surface at North and South latitudes of 61.9º and 51.0º, but only their South latitudes have a path that allows water to transport completely around the surface of the Earth just north of Antarctica .

Table 2: Earth–Moon QCM equilibrium states.

ℓ	$\mathbf{r}_{eq}$ (×10^6 m)	**Latitude**
1	0.143	88.7º
2	0.429	86.1º
3	0.858	82.3º
4	1.430	77.0º
5	2.146	70.3º
6	3.004	61.9º
7	4.005	51.0º
8	5.149	36.1º
9	6.437	—

Note that the $\ell = 1$ to 4 states have equilibrium radii that may be applicable in the Arctic Ocean at the North Pole. The remainder intersect land masses on the surface. All these QCM rotating cylinders could be creating mass currents underneath the crust in the mantle and within interior parts of the Earth, even supporting the generation of the magnetic field and the recent magnetic north pole's rapid movement past the rotational North Pole toward Russia.

Qualitative radial probability distributions for the QCM cylinders that could be affecting the ACC are shown in Figure 2. The verical line at 6.37×10^6 m, is the approximate Earth radius. Their wide radial distributions within the Earth adds to the complexity of interpreting their actions.

As determined below, all displacements from the equilibrium radius will experience a QCM driving force back toward r_{eq}, here interpreted as the distance from the Earth's rotation axis for simplified discussion purposes only. This radial force keeps the ACC roughly localized in the r coordinate, although the qualitative probability distributions shown in Figure 2 reveal a large spread in the radial direction underneath the surface. Moreover, displacements in latitude along the surface are also displacements in the r coordinate, resulting in a complex dynamics to consider in any detailed anaysis.

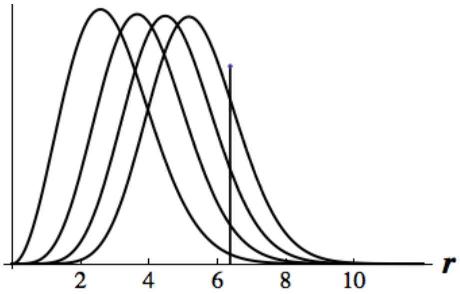


Fig. 2: Representation of the probability distributions for the $\ell = 6$, 7, 8, 9, QCM states with the Earth radius line at 6.37×10^6 m.

A fluid dynamics computer simulation would be needed to better understand the actual behavior of the ACC when QCM forces, winds, Coriolis effects, water density, and water viscosity are accounted for. The atmosphere above the ocean water would also be subject to the QCM forces in both the r direction and the ϕ direction. A rough estimate of the dynamics is considered below.

4.3 Estimates of QCM forces

In the following simplified analysis of the Earth–Moon system, winds and Coriolis forces are ignored. In the ϕ-direction, the QCM angular momentum per unit mass L/μ for a free particle at the equilibrium radius r_{eq} is given by the QCM constraint,

$$\frac{L}{\mu} = m\,c\,H, \tag{7}$$

with $|m| = \ell$ at the Equator, assuming a co-axial cylinder. Thus, substituting $L = \mu\, v_\phi\, r$ for the angular momentum produces the ϕ velocity

$$v_\phi = \frac{mcH}{r}. \tag{8}$$

QCM predicts for the $m = 6$ state a velocity $v_\phi \sim 1.1 \times 10^4$ m/s, and at the $m = 7$ state a $v_\phi \sim 1.26 \times 10^4$ m/s, values which can be compared to the actual average ACC velocity of about 212 m/s with respect to the stars. A large reduction in these predicted ϕ-velocities would be required of viscosity effects in the water and impedance effects of the land interruption at the ocean bottom and at the edges of the continents.

The torque τ required to keep the volume flow V $\sim 4 \times 10^{26}$ m^3/s of ACC moving at approximately 2 m/s faster than the Earth's rotational veleocity can be estimated, using the viscosity $\eta = 1.6$ cP for water at about 2ºC, to be

$$\tau = 2\pi\eta V \approx 10^{14}\ \mathtt{Nm}, \tag{9}$$

which translates to a force of about 10^{-12} N to keep a kilogram of water moving at 2 m/s faster than the Earth's velocity.

Depending upon just where vertically and horizontally one calculates the driving torque pushing the water, QCM force values up to only about 10^{-9} N are estimated to be required. Any vertical water movement at the ACC latitude introduces displacement components in both the radial direction from the rotation axis and in the latitudinal direction. For simplicity, any latitudinal direction movement is assumed to be included within the r direction movement for the spherical geometry of the Earth's surface.

A small displacement from r_{eq} in the radial direction results in an acceleration, calculated by taking the negative gradient of V$_{eff}$, to get

$$a_r = -\frac{r_g c^2}{2r^2} + \frac{m^2 H^2 c^2}{r^3}. \tag{10}$$

So, if the water is at $r > r_{eq}$ or at $r < r_{eq}$, this QCM acceleration tries to move the water back to r_{eq}.

Water temperature differences are important. The surface water may be at a different temperature than the water below, so their density differences produce vertical mixing. Therefore, any water at the QCM equilibrium radius may move to a different radius value, with the radial velocity v_r resulting in a force in the ϕ direction. From Eq. 8, the ϕ acceleration

$$a_\phi = -\frac{mcH}{r^2} v_r. \tag{11}$$

Using the $m = 6$ and $m = 7$ parameters at the ACC, the expression becomes approximately

$$a_\phi = -0.003\ v_r. \tag{12}$$

So both the sinking water and the rising water will experience a ϕ direction acceleration due to the QCM angular momentum per unit mass constraint, the accelerations depending upon the radial velocity directions and magnitudes.

These QCM forces and accelerations, when considered along with Coriolis forces and other influences, could be simulated on computer to determine their relative importance to the transport of the ACC.

Therefore, the estimated results of these QCM derivations suggest force and acceleration values strong enough to keep the ACC transport moving around the Antarctic continent, meaning that the ACC may be in a gravitational energy state dictated by the QCM quantization of angular momentum per unit mass constraint.

5 Conclusion

I have applied the QCM gravitational wave equation to the rotation of the Earth by utilizing both the total vector angular momentum of the Earth's spin plus the larger value of the average angular momentum of the Moon in orbit. The QCM $\ell = 6$ state at $r_{eq} = 3.0 \times 10^6$ m intersects the Earth's surface at 61.9ºS latitude, and the $\ell = 7$ state at $r_{eq} = 4.0 \times 10^6$ m intersects the Earth's surface at 51.0ºS latitude. Both QCM cylinders intersect the surface in the wide latitude region where the ACC flows faster than the Earth's rotation velocity by about 2 m/s.

The enormous QCM predicted velocity of about 1.1×10^4 m/s with respect to the stars is much larger than the actual ACC velocity of about 212 m/s. Viscosity effects on the water transport at all depths would need to be a significant opposing force to be able to reduce this QCM velocity to its actual velocity. Rough estimates of the pertinent forces suggest values on the order of 10^{-12} N to 10^{-9} N per kilogram are required.

Temperature differences produce mixing, which moves water away from the equilibrium radius measured from the rotation axis, resulting in an acceleration in both the r direction and the ϕ direction. The QCM forces combined with the Coriolis force and other effects make for a complex transport of the ACC. However, a computer simulation that includes the QCM force driving the ACC would be necessary in order to evaluate the atmosphere and ocean behaviors in more detail.

Therefore, the QCM wave equation applied in the familiar Schwarzschild metric suggests that the true energy source for the ACC could be gravitational.

Acknowledgements

The author thanks Sciencegems.com for continuing encouragement and support.

Submitted on February 23, 2021

References

1. Zika J. D. et al. Acceleration of the Antarctic Circumpolar Current by Wind Stress along the Coast of Antarctica. *Journal of Physical Oceanography*, 2013, v. 43 (12), 2772–2784.
2. Whitworth III T. The Antarctic Circumpolar Current. *Oceanus*, 1988, v. 32, 53–58.
3. Thompson A. F. and Garabato, A. C. N. Equilibration of the Antarctic Circumpolar Current by Standing Meanders. *Journal of Physical Oceanography*, 2014, v. 44 (7), 1811–1828.
4. Marynets K. On the modeling of the flow of the Antarctic Circumpolar Current. *Monatshefte für Mathematik*, 2019, v. 188, 561–565.
5. Preston H. G. and Potter F. Exploring Large-scale Gravitational Quantization without $\hbar$ in Planetary Systems, Galaxies, and the Universe. arXiv: gr-qc/0303112.
6. Potter F. Update on Pluto and Its 5 Moons Obeying the Quantization of Angular Momentum per Unit Mass Constraint of Quantum Celestial Mechanics. *Progress in Physics*, v. 12 (1), 2016, 56–58. Online: http://www.ptep-online.com/2016/PP-44-10.PDF.
7. Potter F. Multi-planet Exosystems all obey Orbital Angular Momentum Quantization per Unit Mass predicted by Quantum Celestial Mechanics (QCM). *Progress in Physics*, v. 9 (3), 2013, 60–61. Online: http://www.ptep-online.com/2013/PP-34-17.PDF.
8. Potter F. and Preston H. G.. Quantum Celestial Mechanics: large-scale gravitational quantization states in galaxies and the Universe. *1st Crisis in Cosmology Conference: CCC-I*, Lerner E. J. and Almeida J. B. eds., AIP CP822, 2006, 239–252. .
9. Potter F. and Preston H. G. Cosmological Redshift Interpreted as Gravitational Redshift. *Progress in Physics*, v. 3 (2), 2007, 31–33. Online: http://www.ptep-online.com/2007/PP-09-06.PDF.

Stellar Evolution of High Mass-Loss Stars

Pierre A. Millette

E-mail: pierre.millette@uottawa.ca, Ottawa, Canada

In this paper, we investigate the 37 strongest QSO emission lines of stars of type Q in the catalog of Hewitt & Burbidge [22], as determined by Varshni *et al* [21]. We identify the candidate lines from atomic spectra lines data and determine the estimated T_e range down to the 50% ionic element density for the identified candidate emission lines. This information assists in the classification of Q stars from the 37 QSO dominant emission lines. We use the Hertzsprung-Russell diagram to analyze and determine the role of mass loss in the evolution of stars. We review the nucleosynthesis process that leads from massive O stars to WN and then WC Wolf-Rayet stars as a result of mass loss, and then consider the nucleosynthesis of oxygen in massive stars, showing that ^{16}O oxygen has a significant presence in massive stars beyond the WC stage, until the generation of ^{28}Si, where it disappears. This leads us to postulate more than one population of stars of type Q. One group identified by Varshni [27] with O VI and He II emission lines in their spectra implying much higher temperatures and positioning those QSO stars above the WR region in the HRD. The other group with emission lines dominated by temperatures in the O II and O III range, indicating a lower temperature range of QSO stars with a significant number of ionized oxygen lines and some Si emission lines, in addition to the nitrogen WN and carbon WC lines. We postulate that these QSO spectra correspond to unrecognized Wolf-Rayet stars, in particular WO stars and WSi pre-supernova stars, extending into lower temperatures. In that scenario, Q stars would be the end-state of the Wolf-Rayet evolution process, prior to moving to the supernova state.

1 Introduction

In a recent paper [1], we reviewed the physical process of laser action that is occurring in the stellar atmospheres of stars of type W (Wolf-Rayet) and stars of type Q (QSOs), due to the rapid adiabatic expansion of the stellar atmosphere of these stars, resulting in population inversions in the ionic energy levels due to free electron-ion recombination in the cooling plasma. Laser action in non-LTE stellar atmospheres was first proposed by Menzel in [2] and plasma lasers by Gudzenko in [3]. This results in the extremely strong broad emission lines observed in the spectra of these stars.

Significant work has been performed over the last forty years on the analysis of the emission line spectra of WR stars to understand their classification and evolution [4–15] and is still an ongoing area of research. Previously [1], we noted that a similar effort will be required to understand the classification and evolution of stars of type Q, as has been achieved for Wolf-Rayet stars. In this paper, we take an initial stab at this problem.

In addition, we examine the stellar evolution of high mass-loss stars, characterized by WR and QSO stars, and refine our proposal to enhance the Hertzsprung-Russell Diagram (HRD) by including stars of type W and stars of type Q in the HRD [1]. This allows us to postulate that QSO stars can be identified as unrecognized Wolf-Rayet stars, in particular WO stars and WSi pre-supernova stars, to position them on the HRD, and provide a better understanding of the evolution of high mass-loss stars as displayed in the Hertzsprung-Russell Diagram.

2 Spectra of stars of type W

The emission line spectra of Wolf-Rayet stars are dominated by lines of helium He, carbon C, nitrogen N and oxygen O. The spectra fall into two broad classes: WN stars, which have prominent lines of nitrogen N and helium He ions, with a very strong He II Pickering series ($n = 4 \to n'$), and essentially no lines of carbon C; and WC stars, where the lines of carbon C and oxygen O are prominent along with the helium He ions, while those of nitrogen N seem to be practically absent [19, p. 485]. An additional subtype WO with strong O VI lines has also recently been added as a separate subtype. The spectra are characterized by the dominance of emission lines, notable for the almost total absence of hydrogen H lines [4].

The number of WR stars in our galaxy is small: the 2001 VII$^{\text{th}}$ catalog of galactic WR stars gave the number at 227 stars, comprised of 127 WN stars, 87 WC stars, 10 WN/WC stars and 3 WO stars [16]. A 2006 update added another 72 WR stars, including 45 WN stars, 26 WC stars and one WO star [17]. The latest number from the August 2020 Galactic Wolf-Rayet Catalogue v1.25 is 667 WR stars [18].

Wolf-Rayet stars have extended atmospheres whose thickness is an appreciable fraction of their stellar radius [19, p. 243]. The material generating the lines is flowing outward from the stellar photosphere. These flows are driven by radiation pressure acting on the stellar atmosphere. Mass loss in stellar winds, particularly in WR stars, is well established [19, pp. 266, 523]. Mass loss rates $\mathcal{M}$ for WR stars are estimated to be of order 10^{-5} up to perhaps 10^{-4} $\mathcal{M}_\odot$/year [20, p. 628] — for comparison, the mass loss rate for the solar wind is about $10^{-14}\ \mathcal{M}_\odot$/year. The massive trans-sonic stellar winds flow velocities in WR stars rise from close to zero in the stellar photosphere to highly supersonic values of order 3 000 km/s within one stellar radius from the surface. Rapid cooling of the strongly ionized plasma results in rapid recombination of the free electrons and the ions into highly excited ionic states, resulting in population inversions and laser action.

3 Spectra of stars of type Q

Similar physical processes are expected to predominate in QSO stars due to the physical similarity of WR and QSO spectra, including the almost total absence of hydrogen H lines. However, while the number of identified WR stars is relatively small, the number of identified QSO stars is larger as we will see later.

Varshni *et al* [21] studied the distribution of QSO spectral emission lines (in the observed frame, i.e. unshifted) of 7 315 QSOs from the catalog of Hewitt & Burbidge [22], of which 5 176 have emission lines. This resulted in a total of 14 277 emission lines in the range λ1271 to λ17993, with the vast majority in the visual range λ3200 to λ5600*.

A number of very strong peaks were found in a histogram of the statistical frequency of the emission lines against wavelength using 4 Å bins. The emission line distribution was expressed in units of standard deviation above a random average, to ensure the lines are statistically significant. The 37 strongest QSO emission lines in the catalog were more than four standard deviations above the random average, of which 13 peaks were above 5σ, of which 3 peaks were above 6σ, of which one peak was above 8σ. These lines are given in Table 1 including the number of standard deviations above a random average.

The 37 QSO emission lines were compared with Wolf-Rayet emission lines in [21], and 27 were found to also occur in WR star spectra, 7 in novae-like star spectra, and two in novae spectra. These are also included in Table 1, along with corresponding candidate element emission lines identification. The lines have been compared against existing sources of data such as Willett's [23] and Bennett's [24] lists of laser transitions observed in laboratories and the *NIST Atomic Spectra Database Lines Data* [25].

In Table 2, we have added the estimated T_e range down to the 50% ionic element density obtained from House [26] for the identified candidate element emission lines. In Tables 3 and 4, we determine the best known line identification and estimated T_e range down to the 50% ionic element density from House [26] respectively, for the 37 QSO emission lines identified in [21]. This provides information to assist in the classification of Q stars from the 37 QSO dominant emission lines.

Varshni [27] also investigated O VI and He II emission lines in the spectra of QSOs, planetary nebulae and Sanduleak stars (WO stars characterized by a strong O VI emission line at λ3811.34 — one example, blue supergiant star Sk -69 202, was identified as the progenitor of supernova 1987A). O VI emission lines imply much higher temperatures (180 000 K < O VI < 230 000 K) than those of Table 4, which are dominated by temperatures in the O II and O III range (16 000 K < O II < 46 000 K and 46 000 K < O III < 73 000 K respectively).

4 Comparative numbers of O, W amd Q stars

The comparative numbers of O, W and Q stars provide hints on their relative classification with respect to their evolution and the Hertzsprung-Russell diagram. O stars are known to be massive hot blue-white stars with surface temperatures in excess of 30 000 K. Wolf-Rayet stars have typical masses in the range of 10–25 $M_\odot$, extending up to 80 $M_\odot$ for hydrogen-rich WN stars [5], and surface temperatures ranging from 30 000 K to around 210 000 K.

Conti *et al* [6] measured the actual numbers and distributions of O stars and Wolf-Rayet stars in a volume-limited sample of stars within 2.5 kpc of the Sun. They found the observed WR/O star number ratio to be given by

$$\frac{\mathrm{WR}}{\mathrm{O}\,(M \geq 40 \pm 5M_\odot)} = 0.36 \pm 0.15\,. \tag{1}$$

The distribution of WR stars matches that of massive O stars, primarily close to the galactic plane, predominantly in spiral arms Population I stars, which is seen to indicate that WR stars are descendant from the most luminous and massive O stars, likely due to mass loss.

From the latest number of 667 WR stars seen previously in §2, we obtain an estimated number of O stars of

$$1\,850\,^{+1\,350}_{-550}\,,$$

that is between 1 300 and 3 200, from (1). These results are in the same ballpark as available catalogues of O stars, which are still very much a work in progress [28–33]. This number is similar to that of planetary nebulae with about 2 700 known in 2008 (MASH catalogue) [34, 35]. These distributions and numbers of O stars and Wolf-Rayet stars agree with the changes to the Hertzsprung-Russell diagram suggested in [1] to include more massive and hotter stars of type W beyond the stars of type O B.

For the number of stars of type Q, we saw previously in §3 that the Hewitt & Burbidge catalog of 1993 [22] included 7 315 QSOs, of which 5 176 have emission lines, which represents 11 times the current number of known WR stars. However, the latest edition of the *Sloan Digital Sky Survey Quasar Catalog DR16Q* [36] to August 2018, includes a total of 750 414 quasars (100 times the 1993 Hewitt & Burbidge catalog number). This represents 1 125 times the number of WR

---

*The notation λ indicates wavelengths measured in Å.

QSO λ (Å)	σ	WR λ (Å)	NL λ (Å)	Nova λ (Å)	Emitter λ candidates	Emitter λ candidates	Emitter λ candidates
3356	4.0	3358.6			N III λ3355	O III λ3355.9	C III λ3358
3489	4.1	3493			O IV λ3489.83	O II λ3488.258	O II λ3494.04
3526	6.7				C V λ3526.665	O II λ3525.567	
3549	4.3				O II λ3549.091	Si III λ3549.42	
3610	4.6	⟨3609.5⟩			C III λ3609.6	He I λ3613.6	N II λ3609.097
3648	5.4	3645.4			O II λ3646.56	O IV λ3647.53	O IV λ3642
3683	4.4	3687	3685.10		O II λ3683.326	O III λ3682.383/6.393	N IV λ3689.94
3719	4.7	⟨3721.0⟩			O III λ3721	O V λ3717.31	
3770	4.7	⟨3769.5⟩			N III λ3770.36/1.05	Si III λ3770.585	O III λ3774.026
3781	5.3	3784.8			He II λ3781.68	O II λ3784.98	N III λ3779
3831	5.1	3829.9			He II λ3833.80	C II λ3831.726	O II λ3830.29
3842	4.7				O IV λ3841.07	N II λ3842.187/.449	O II λ3842.815
3855	5.0	3856.6			N II λ3855.096/.374	O II λ3856.134	He II λ3858.07
3890	8.4	⟨3889.0⟩			He I λ3888.64	C III λ3889.18/.670	O III λ3891.759
3903	5.6		3903.0		O III λ3903.044		
3952	4.4	⟨3954.2⟩			O II λ3954.3619	Si III λ3952.23/3.071	
4012	6.0	⟨4008.4⟩			N III λ4007.88	N III λ4013.00	Si III λ4010.236
4135	4.8	?			N III λ4134.91/6.07	O V λ4134.11	N II λ4133.673
4276	5.9	?	4276.6	4275.5	O II λ4275.5	O II λ4275.994	O II λ4276.620
4524	4.7	⟨4520.4⟩			N III λ4523.56/7.9	O III λ4524.2/7.3	O V λ4522.66
4647	4.0	4650.8			C III λ4647.40/51.35	O II λ4647.803/9.1348	O III λ4649.973
4693	4.7	?	4697.0		O II λ4693.195	N II λ4694.274/7.638	O III λ4696.225
4771	4.1	?	4772.1		N IV λ4769.86	O IV λ4772.6	O II λ4773.782
4801	4.3	⟨4799.6⟩			O IV λ4800.74	Si III λ4800.43	
4817	4.4	4814.6	4814.4		O IV λ4813.15	Si III λ4813.33/9.72	N II λ4815.617
4910	4.9	4909.2			N III λ4904.78	Si III λ4912.310	
4925	4.5	⟨4924.6⟩			O II λ4924.531	He I λ4921.9	
4956	7.0	4958		4959.0	O II λ4955.705	O III λ4958.911	
5018	5.6	⟨5018.3⟩			He I λ5015.67	C IV λ5015.9/7.7	N II λ5016.39
5035	4.2				N III λ5038.31		
5049	5.5	5049.9			He I λ5047.7	C III λ5048.95	O III λ5049.870
5096	5.5	5092.9			N III λ5097.24	O III λ5091.880	O II λ5090.920
5111	4.8		5111.5		O II λ5110.300/1.913	Si III λ5111.1	O III λ5112.18
5173	4.6	5171.1			N II λ5171.266/2.344	N II λ5173.385	O III λ5171.29
5266	5.3	5266.3			O III λ5268.301		
5345	4.1	5343.3			O II λ5344.104	C III λ5345.881	
5466	4.0	⟨5469.9⟩			Si II λ5466.43/9.21	O V λ5471.12	Si III λ5473.05

Table 1: QSO emission lines in the range λ3200 to λ5600 from Varshni *et al* [21] (NL: novae-like star).

QSO λ (Å)	σ	WR λ (Å)	NL λ (Å)	Nova λ (Å)	Emitter T_e (K) candidates	Emitter T_e (K) candidates	Emitter T_e (K) candidates
3356	4.0	3358.6			33k < N III < 65k	46k < O III < 73k	29k < C III < 58k
3489	4.1	3493			73k < O IV < 130k	16k < O II < 46k	16k < O II < 46k
3526	6.7				92k < C V < 730k	16k < O II < 46k	
3549	4.3				16k < O II < 46k	18k < Si III < 46k	
3610	4.6	⟨3609.5⟩			29k < C III < 58k	He I < 26k	18k < N II < 33k
3648	5.4	3645.4			16k < O II < 46k	73k < O IV < 130k	73k < O IV < 130k
3683	4.4	3687	3685.10		16k < O II < 46k	46k < O III < 73k	65k < N IV < 103k
3719	4.7	⟨3721.0⟩			46k < O III < 73k	146k < O V < 184k	
3770	4.7	⟨3769.5⟩			33k < N III < 65k	18k < Si III < 46k	46k < O III < 73k
3781	5.3	3784.8			26k < He II < 73k	16k < O II < 46k	33k < N III < 65k
3831	5.1	3829.9			26k < He II < 73k	15k < C II < 29k	16k < O II < 46k
3842	4.7				73k < O IV < 130k	18k < N II < 33k	16k < O II < 46k
3855	5.0	3856.6			18k < N II < 33k	16k < O II < 46k	26k < He II < 73k
3890	8.4	⟨3889.0⟩			He I < 26k	29k < C III < 58k	46k < O III < 73k
3903	5.6		3903.0		46k < O III < 73k		
3952	4.4	⟨3954.2⟩			16k < O II < 46k	18k < Si III < 46k	
4012	6.0	⟨4008.4⟩			33k < N III < 65k	18k < Si III < 46k	
4135	4.8	?			33k < N III < 65k	146k < O V < 184k	18k < N II < 33k
4276	5.9	?	4276.6	4275.5	16k < O II < 46k		
4524	4.7	⟨4520.4⟩			33k < N III < 65k	46k < O III < 73k	146k < O V < 184k
4647	4.0	4650.8			29k < C III < 58k	16k < O II < 46k	46k < O III < 73k
4693	4.7	?	4697.0		16k < O II < 46k	18k < N II < 33k	46k < O III < 73k
4771	4.1	?	4772.1		65k < N IV < 103k	73k < O IV < 130k	16k < O II < 46k
4801	4.3	⟨4799.6⟩			73k < O IV < 130k	18k < Si III < 46k	
4817	4.4	4814.6	4814.4		73k < O IV < 130k	18k < Si III < 46k	18k < N II < 33k
4910	4.9	4909.2			33k < N III < 65k	18k < Si III < 46k	
4925	4.5	⟨4924.6⟩			16k < O II < 46k	He I < 26k	
4956	7.0	4958		4959.0	16k < O II < 46k	46k < O III < 73k	
5018	5.6	⟨5018.3⟩			He I < 26k	65k < C IV < 92k	18k < N II < 33k
5035	4.2				33k < N III < 65k		
5049	5.5	5049.9			He I < 26k	29k < C III < 58k	46k < O III < 73k
5096	5.5	5092.9			33k < N III < 65k	46k < O III < 73k	16k < O II < 46k
5111	4.8		5111.5		16k < O II < 46k	18k < Si III < 46k	46k < O III < 73k
5173	4.6	5171.1			18k < N II < 33k	46k < O III < 73k	
5266	5.3	5266.3			46k < O III < 73k		
5345	4.1	5343.3			16k < O II < 46k	29k < C III < 58k	
5466	4.0	⟨5469.9⟩			8.2k < Si II < 18k	146k < O V < 184k	18k < Si III < 46k

Table 2: QSO emission lines in the range λ3200 to λ5600 with estimated T_e range down to the 50% ionic element density from House [26].

QSO λ (Å)	WR λ (Å)	NL λ (Å)	Nova λ (Å)	Emitter QSO λ	Emitter WR λ	Emitter NL λ	Emitter Nova λ
3356	3358.6			O III λ3355.9	C III λ3358		
3489	3493			O II λ3488.258	O II λ3494.04		
3526				O II λ3525.567			
3549				O II λ3549.091			
3610	⟨3609.5⟩			C III λ3609.6			
3648	3645.4			O IV λ3647.53	O II λ3646.56		
3683	3687	3685.10		O II λ3683.326	O III λ3686.393	O III λ3686.393	
3719	⟨3721.0⟩			O III λ3721	O III λ3721		
3770	⟨3769.5⟩			N III λ3770.36	N III λ3770.36		
3781	3784.8			He II λ3781.68	O II λ3784.98		
3831	3829.9			C II λ3831.726	O II λ3830.29		
3842				N II λ3842.187			
3855	3856.6			N II λ3855.096	O II λ3856.134		
3890	⟨3889.0⟩			C III λ3889.670	C III λ3889.18		
3903		3903.0		O III λ3903.044		O III λ3903.044	
3952	⟨3954.2⟩			Si III λ3952.23	O II λ3954.3619		
4012	⟨4008.4⟩			N III λ4013.00	N III λ4007.88		
4135	?			N III λ4134.91			
4276	?	4276.6	4275.5	O II λ4275.994		O II λ4276.620	O II λ4275.5
4524	⟨4520.4⟩			O III λ4524.2	O V λ4522.66		
4647	4650.8			C III λ4647.40	C III λ4651.35		
4693	?	4697.0		O II λ4693.195		O III λ4696.225	
4771	?	4772.1		N IV λ4769.86		O IV λ4772.6	
4801	⟨4799.6⟩			O IV λ4800.74	O IV λ4800.74		
4817	4814.6	4814.4		N II λ4815.617	Si III λ4813.33	O IV λ4813.15	
4910	4909.2			Si III λ4912.310	Si III λ4912.310		
4925	⟨4924.6⟩			O II λ4924.531	O II λ4924.531		
4956	4958		4959.0	O II λ4955.705	O III λ4958.911		O III λ4958.911
5018	⟨5018.3⟩			C IV λ5017.7	C IV λ5017.7		
5035				N III λ5038.31?			
5049	5049.9			C III λ5048.95	O III λ5049.870		
5096	5092.9			N III λ5097.24	O III λ5091.880		
5111		5111.5		Si III λ5111.1		O II λ5111.913	
5173	5171.1			N II λ5173.385	N II λ5171.266		
5266	5266.3			O III λ5268.301?	O III λ5268.301?		
5345	5343.3			O II λ5344.104	O II λ5344.104		
5466	⟨5469.9⟩			Si II λ5466.43	Si II λ5469.21		

Table 3: QSO emission lines in the range λ3200 to λ5600 with best known line identification.

QSO λ (Å)	WR λ (Å)	NL λ (Å)	Nova λ (Å)	Emitter QSO T_e	Emitter WR T_e	Emitter NL T_e	Emitter Nova T_e
3356	3358.6			46k<O III<73k	29k<C III<58k		
3489	3493			16k<O II<46k	16k<O II<46k		
3526				16k<O II<46k			
3549				16k<O II<46k			
3610	⟨3609.5⟩			29k<C III<58k			
3648	3645.4			73k<O IV<130k	16k<O II<46k		
3683	3687	3685.10		16k<O II<46k	46k<O III<73k	46k<O III<73k	
3719	⟨3721.0⟩			46k<O III<73k	46k<O III<73k		
3770	⟨3769.5⟩			33k<N III<65k	33k<N III<65k		
3781	3784.8			26k<He II<73k	16k<O II<46k		
3831	3829.9			15k<C II<29k	16k<O II<46k		
3842				18k<N II<33k			
3855	3856.6			18k<N II<33k	16k<O II<46k		
3890	⟨3889.0⟩			29k<C III<58k	29k<C III<58k		
3903		3903.0		46k<O III<73k		46k<O III<73k	
3952	⟨3954.2⟩			18k<Si III<46k	16k<O II<46k		
4012	⟨4008.4⟩			33k<N III<65k	33k<N III<65k		
4135	?			33k<N III<65k			
4276	?	4276.6	4275.5	16k<O II<46k		16k<O II<46k	16k<O II<46k
4524	⟨4520.4⟩			46k<O III<73k	146k<O V<184k		
4647	4650.8			29k<C III<58k	29k<C III<58k		
4693	?	4697.0		16k<O II<46k		46k<O III<73k	
4771	?	4772.1		65k<N IV<103k		73k<O IV<130k	
4801	⟨4799.6⟩			73k<O IV<130k	73k<O IV<130k		
4817	4814.6	4814.4		18k<N II<33k	18k<Si III<46k	73k<O IV<130k	
4910	4909.2			18k<Si III<46k	18k<Si III<46k		
4925	⟨4924.6⟩			16k<O II<46k	16k<O II<46k		
4956	4958		4959.0	16k<O II<46k	46k<O III<73k		46k<O III<73k
5018	⟨5018.3⟩			65k<C IV<92k	65k<C IV<92k		
5035				33k<N III<65k			
5049	5049.9			29k<C III<58k	46k<O III<73k		
5096	5092.9			33k<N III<65k	46k<O III<73k		
5111		5111.5		18k<Si III<46k		16k<O II<46k	
5173	5171.1			18k<N II<33k	18k<N II<33k		
5266	5266.3			46k<O III<73k	46k<O III<73k		
5345	5343.3			16k<O II<46k	16k<O II<46k		
5466	⟨5469.9⟩			8.2k<Si II<18k	8.2k<Si II<18k		

Table 4: QSO emission lines in the range λ3200 to λ5600 with best known line identification and with estimated T_e range down to the 50% ionic element density from House [26].

stars — quite obviously, we are dealing with a phenomenon that is not as rare as O stars, WR stars or planetary nebulae.

There is a total of between 100 billion and 400 billion stars estimated in the Milky Way galaxy. The number of QSOs just represents about 2×10^{-6} times the estimated number of stars in the Milky Way galaxy, thus a fairly rare phenomenon, even if it is about 10^3 times more numerous than Wolf-Rayet stars. This requires further analysis in terms of understanding and positioning stars of type Q on the Hertzsprung-Russell diagram, with some of the QSOs with strong O VI lines [27] likely lying beyond the stars of type W O B as suggested in [1] and some having temperatures in the WR star range. The estimated number of quasars may be inflated due to the tendency in modern astronomy to identify redshifts as a predominant causative factor, but even if it is off by a factor of ten, the QSO phenomenon is much more common than Wolf-Rayet stars, in spite of their similarity.

We take a brief look at novae and novae-like stars [37], given their presence in Tables 1 to 4. These are part of what are known as *cataclysmic variables* (CVs), which are binary star systems consisting of a white dwarf and a normal star companion. Matter transfer to the white dwarf from the companion star results in the formation of an accretion disk around the white dwarf, which produces occasional cataclysmic outbursts of matter.

A main sequence star in a binary system evolves into a white dwarf for a mass below the Chandrasekhar limit (white dwarf maximum mass limit of about $1.4\,M_\odot$). Novae are CV white dwarfs that undergo an eruption that can change by 10–12 magnitudes in a few hours. They are subdivided into classical novae (single observed eruption with a spectroscopically detected shell of ejected matter), recurrent novae (multiple observed outbursts with detected shell of matter), and dwarf nova (multiple observed eruptions with no shell of detected matter).

Nova-like (NL) variables include all “non-eruptive” cataclysmic variables. These systems have spectra, mostly emission spectra, indicating that they are possibly novae that have not been observed. A catalogue of cataclysmic variables to 2006 contains 1 600 CVs [38].

5 The Hertzsprung-Russell diagram and the role of mass loss in the evolution of stars

The Hertzsprung-Russell diagram is a powerful tool to analyze and represent stellar evolution and understand the characteristics and properties of stars. Most HR diagrams cover the temperature range 40 000 K and below, thus ignoring hotter and more massive stars of interest in this work.

From an idealized perspective, the main sequence is a vaguely diagonal curve running from the upper left to the lower central part of the diagram; from it, vaguely horizontal branches tend to the right of the diagram. The main sequence is known as the Zero-Age Main Sequence (ZAMS) which a star enters when it starts core hydrogen burning; massive stars (O,B) rapidly burn the hydrogen in $\sim 3 \times 10^6$ years, while low mass stars (M) burn the hydrogen more slowly in $\sim 2 \times 10^{11}$ years.

As the core hydrogen becomes depleted, the star moves towards the horizontal portion of the diagram, and once core hydrogen burning terminates, it moves towards the right on the horizontal branch, becoming a red giant for cooler less massive stars (G) or a red supergiant for hot massive stars (O). Interestingly enough, in a recent study of stars of types O and early-B in the Wing of the Small Magellanic Cloud (SMC) satellite galaxy, Ramachandra *et al* [39] have found that the above scenario applies to O stars with initial mass below $\sim 30\,M_\odot$, while O stars with initial mass above $\sim 30\,M_\odot$ appear to always stay hot.

Once a star has exhausted its core hydrogen (and hydrogen shell), it enters its core helium burning phase. In Fig. 1, we reproduce the very important Hertzsprung-Russell diagram of [39] for the stars of types O and early-B of the SMC Wing: it covers the temperature range up to 200 000 K, shows the Helium Zero-Age Main Sequence (He-ZAMS) and also exceptionally includes the Wolf-Rayet (WR) stars. As we saw previously, the red giant and supergiant scenario, where the hydrogen-depleted stars veer off the ZAMS to the right, applies to O stars with initial mass below $\sim 30\,M_\odot$ (shaded portion in Fig. 1). However, as we see in Fig. 1, for O stars with initial mass above $\sim 30\,M_\odot$, the hydrogen-depleted stars veer off the ZAMS to the left to become WR stars, which are known to be hydrogen-deficient. Of the main factors affecting massive star evolution, focusing on rotation, binarity and mass-loss rate, we believe this dichotomy in behaviour is because of the massive mass loss in WR stars as seen in §2 driving laser action in their stellar atmospheres, while [39] believes it is due to the rapid rotation of the stars, leading to efficient mixing of the stellar interior and quasi-chemically homogeneous evolution (QCHE) .

The work of Ramachandra *et al* [39] is an excellent example of using one of the better tools at our disposal to understand stellar astrophysical problems by performing analysis on the observed data in neighbouring galaxies, as mentioned in [1]. Along those lines, Hainich *et al* [40] has performed an analysis of single WN Wolf-Rayet stars in the Small Magellanic Cloud. Fig. 2 is a reproduction of the Hertzsprung-Russell diagram for the WN stars of the Small Magellanic Cloud (SMC) from Hainich *et al* [40], which also includes the WN stars of the Large Magellanic Cloud (LMC) and the Milky Way (MW). It corresponds to the upper left portion of Ramachandra’s HRD for log luminosity > 5.2 and temperatures $> 25\,000$ K (in the WR region), and provides details for the WNE and WNL populations of the SMC, LMC and Milky Way galaxy. WNE is a subtype for early-type WN stars (WN2–WN5), while WNL is for late-type WN stars (WN6–WN11).

This HRD provides more details on WN star properties:

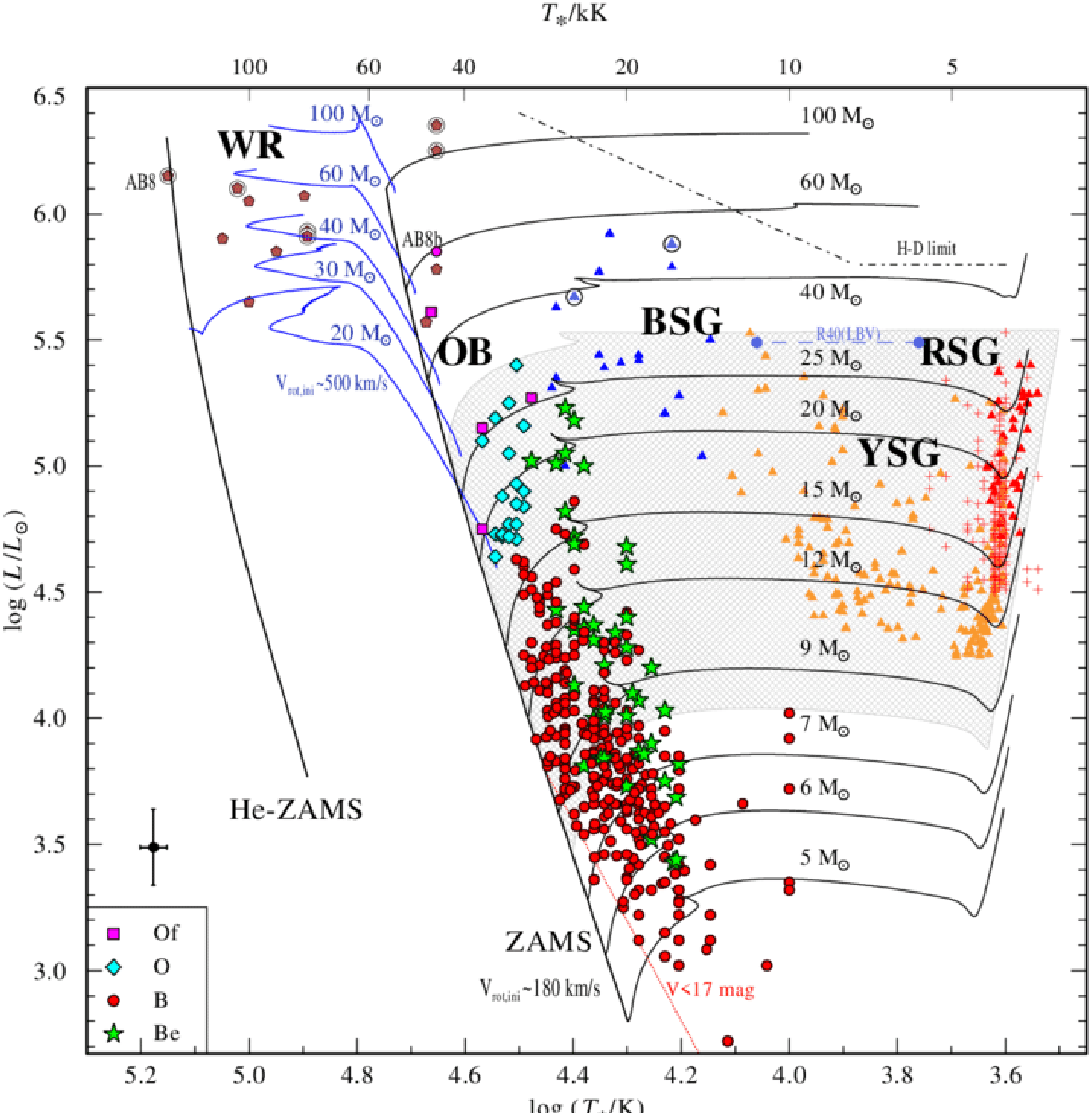


Fig. 1: Hertzsprung Russell diagram for massive stars in the Wing of the SMC reproduced from Ramachandra *et al* [39]. Typical error bar shown at bottom left corner. The brown pentagons represent WR stars (encircled if in a binary system), yellow symbols represent yellow supergiant (YSG) stars, blue triangles for BSGs (blue supergiant), and red triangles for RSGs (red supergiant). Black tracks show standard stellar evolutionary paths, while the blue tracks show the paths of quasi-chemically homogeneously evolving (QCHE) WR stars.

most WNE stars are on the left of the ZAMS line, but to the right of the He-ZAMS line; while most WNL stars are on the right of the ZAMS line, but close to it in the hydrogen depletion region of the stellar evolution curve, above log luminosity > 5.5 corresponding to stellar masses where the stars do not evolve into colder supergiants, as mentioned by Ramachandra *et al* [39]. See also Figures 7 and 8 of [11] for WN stars in the Large Magellanic Cloud. Thus, the calculated WR star evolution curves that extend to the right into lower temperature supergiant stars, usually seen in published HR diagrams, are likely incorrect, especially considering their high mass-loss rates driving laser action in their stellar atmospheres. The Luminous Blue Variable (LBV) stars included in such stellar evolution curves are more than likely variable Wolf-Rayet stars.

Metallicity is a measure of the abundance of elements heavier than hydrogen or helium in an astronomical object. Hence stars and nebulae with relatively high carbon, nitrogen, oxygen, neon, *etc* abundance have high metallicity values z (the metallicity of the Sun is $z = 0.0134$). The degree of wind mass-loss of WR stars depends on their initial metallicity. Metallicity thus has an effect on the evolution of massive stars and of WR stars in particular. The Small Magellanic Cloud is a low-metallicity environment, lower than the metal-

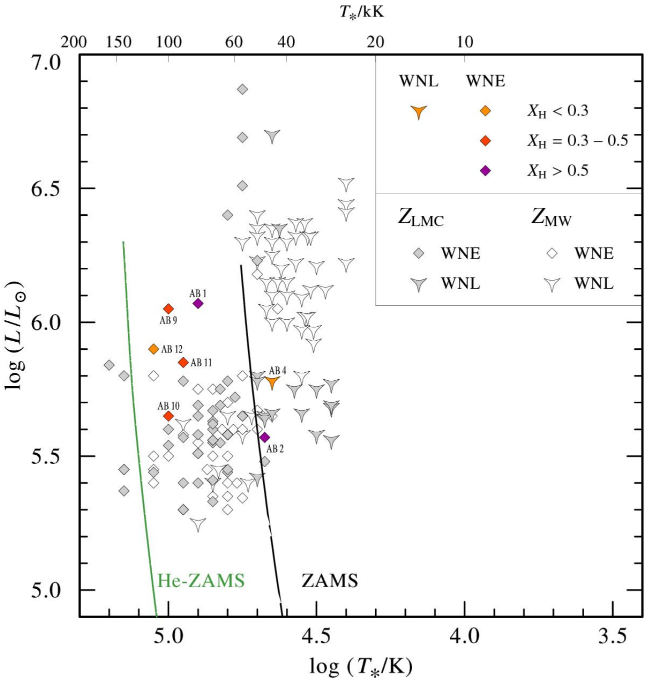


Fig. 2: Hertzsprung-Russell diagram for WN stars in the Small Magellanic Cloud (SMC) reproduced from Hainich *et al* [40]. Includes the Large Magellanic Cloud (LMC gray-filled symbols) and Milky Way (MW open symbols) WN stars.

deficient Large Magellanic Cloud itself lower in comparison to the Milky Way. SMC WN stars thus have on average lower mass-loss rates and weaker winds than their counterparts in the LMC and the Milky Way [40]. A reduction in the mass-loss rate at lower metallicity results in weaker emission line spectra in WR stars, a clear indication that the strong emission lines are due to the mass-loss rate which results in lasing transitions as seen in [1].

The process that leads from massive O stars to WR stars as a result of mass loss is believed to be well understood [41]. As a massive star evolves and loses mass, it eventually exposes He and N (the products of CNO burning) at the surface and is then spectroscopically identified as a WN star. As the star continues losing mass, it eventually exposes C and O (the products of He burning) at the surface and is then identified as a WC star. The mass loss rates depend on the metallicity of the environment which results in different WC/WN ratios as observed in Local Group galaxies. This is reflected in lower WC/WN ratios in lower metallicity environments: $(\mathrm{WC/WN})_{\mathrm{SMC}} = 0$, $(\mathrm{WC/WN})_{\mathrm{LMC}} = 0.25$ and $(\mathrm{WC/WN})_{\mathrm{MW}} = 1$ [40].

Meyer *et al* [42] have analyzed the nucleosynthesis of oxygen in massive stars (see also [43]). In their model calculations, they find that in the WC stars, the oxygen in the C/O zone is dominated by the ^{16}O isotope. This matter which is part of the helium burning core, does not partake in the carbon shell burning. This is followed by the O/Ne zone where the star experiences convective shell carbon burning and where there is a slight ^{16}O depletion, but where ^{16}O still strongly dominates the oxygen abundances. This is followed by the O/Si zone where the star experiences shell neon burning which increases the ^{16}O slightly. Finally, the star burns its ^{16}O into ^{28}Si and heavier isotopes both in pre-supernova and supernova nucleosynthesis, devoid of any oxygen.

Thus they find that oxygen has a significant presence in massive stars beyond the WC stage, until the generation of ^{28}Si, where it disappears. Considering Table 3, this behavior is interesting due to the presence of, in addition to the ionized nitrogen and carbon lines, a significant number of ionized oxygen lines, and the presence of some standalone silicon lines.

6 The evolution of stars of type Q

Given all of these considerations, how does the evolution of stars of type Q fit in the Hertzsprung-Russell diagram? We know that they are undergoing high mass-loss due to the broad high intensity spectral lines indicative of laser action in their stellar atmosphere. As seen in Table 3, their emission spectra are dominated by lines of ionized He, C, N, O and Si, with many lines in common with WR stars and novae-like stars.

There may be more than one population of stars of type Q. One group identified by Varshni [27] with O VI and He II emission lines in their spectra, in common with planetary nebulae and Sanduleak stars, implies much higher temperatures in the range 180 000 K < O VI < 230 000 K, positioning those QSO stars above the WR region in the HRD of Ramachandra *et al* [39] given in Fig. 1. However, there are emission lines as given in Table 3, which are dominated by temperatures in the O II and O III range (16 000 K < O II < 46 000 K and 46 000 K < O III < 73 000 K respectively), indicating a lower temperature range of QSO stars.

Indeed, as seen previously, these QSO emission line spectra have a significant number of ionized oxygen lines. WN and WC Wolf-Rayet stars predominate, with WN stars having the upper hand in low metallicity environments. However, the recently recognized WO lines are rare — could the QSO spectra with a significant number of ionized O emission lines and some Si emission lines, correspond to unrecognized much more numerous WO Wolf-Rayet stars extending into lower temperatures? They would in effect fill up the Hertzsprung-Russell diagram of Ramachandra *et al* [39] given in Fig. 1 in the range $16\,000\,\mathrm{K} < T_e < 73\,000\,\mathrm{K}$ for stellar masses above $\sim 30\,M_\odot$.

For example, if we look at QSO 3C 273, the first radio source quasar for which an optical counterpart was identified in 1963, its spectrum consisted of one strong emission line and one medium to weak strength line (λ5637, λ7588). Comparing these lines against existing sources of data [23–25], the following identifications are obtained from the *NIST*

Atomic Spectra Database Lines Data and the corresponding estimated T_e range down to the 50% ionic element density obtained from House [26] for the identified candidate element emission lines (see Table 5). Based on this information, we would be inclined to conclude that the broad observed emission lines correspond to C II λ5640.55 and O II λ7593, with an estimated stellar temperature in the range 16 000 K < T_e < 29 000 K.

QSO λ (Å)	strength	Emitter λ	Emitter T_e
5637	S	C II λ5640.55	15k<C II<29k
		Si II λ5633/41	8.2k<Si II<18k
7588	M-W	O II λ7593	16k<O II<46k
		C III λ7586.41	29k<C III<58k
		O IV λ7585.74	73k<O IV<130k

Table 5: QSO 3C 273 observed emission lines, identification and estimated T_e range down to the 50% ionic element density obtained from House [26].

7 Laser action in WR and QSO stars

The details of the process of laser action in the stellar atmospheres of Wolf-Rayet stars and Quasi-Stellar Object stars are given in [1]. The physical process of population inversions in expanding stellar atmospheres led Varshni to formulate his Plasma Laser Star (PLS) model as an explanation of the spectra of Wolf-Rayet stars and Quasi-Stellar Objects [47–52]. Model calculations starting from an initial element number density of 10^{14} cm^{-3} are performed for a grid of free electron number density n_e and temperature T_e values. The population inversion is displayed on $n_e - T_e$ diagrams showing contours of equal P or α', where [44]

$$P = \frac{n_q}{\omega_q} - \frac{n_p}{\omega_p}\,, \tag{2}$$

where n_q is the population density and ω_q is the statistical weight of level q, and [45, p, 23]

$$\alpha = \sqrt{\frac{\ln 2}{\pi}}\left(\frac{\omega_q\, A_{q\to p}}{4\pi}\right)\frac{P\,\lambda_0^2}{\Delta\nu}\,, \tag{3}$$

where λ_0 is the centre wavelength of the transition, $\Delta\nu$ is the linewidth, $A_{q\to p}$ is the Einstein probability coefficient for spontaneous transition from level q to p, and $\alpha' = \alpha\,\Delta\nu$. Fig. 3 shows a typical $n_e - T_e$ diagram with equi-α' contours for inversely populated transition $6f \to 5d$ of C IV.

Taking Quasi-Stellar Objects to be local stellar objects instead of distant galactic objects eliminates the problems associated with their currently accepted cosmological interpretation: energy source, superluminal velocities, optical variability, quasar proper motions [53, 54], quasar binary systems [55, 56], naked (no nebulosity) quasars, *etc*. The properties of QSOs are similar to those of WR stars and, as stars, those are easily explainable in terms of commonly known physical processes. QSO stars could very well be unrecognized Wolf-Rayet stars, in particular WO stars and WSi pre-supernova stars. In that case, Q stars would be the end-state of the Wolf-Rayet evolution process and would account for their number larger than WN and WC W stars (about 10^3 times), prior to moving to the supernova state.

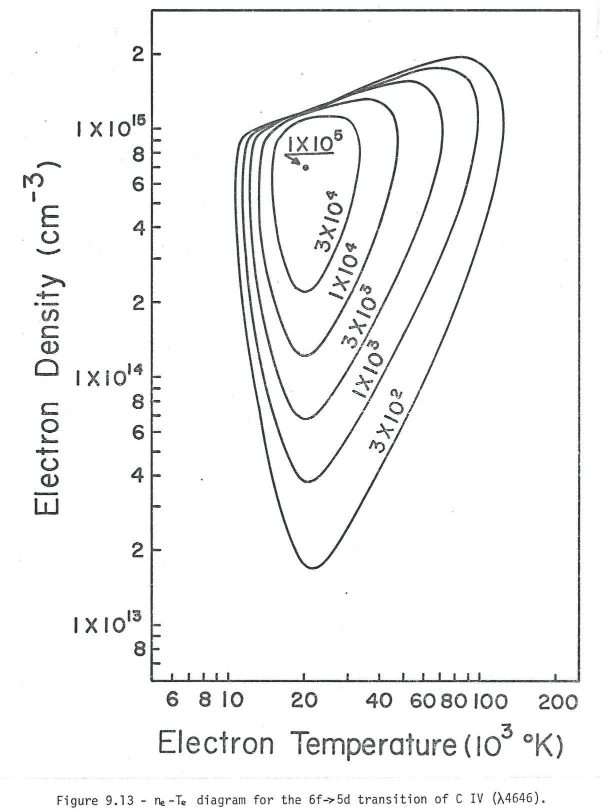


Fig. 3: Typical $n_e - T_e$ diagram showing laser gain equi-α' contours in cm^{-1} s^{-1} for the $6f \to 5d$ transition of C IV [46, p. 257].

8 Discussion and conclusion

In this paper, we have investigated the 37 strongest QSO emission lines of stars of type Q in the catalog of Hewitt & Burbidge [22], investigated by Varshni *et al* [21]. We have used Willett's [23] and Bennett's [24] lists of laser transitions observed in laboratories and the *NIST Atomic Spectra Database Lines Data* [25] to identify candidate lines. In addition, we have determined the estimated T_e range down to the 50% ionic element density obtained from House [26] for the identified candidate element emission lines. This information assists in the classification of Q stars from the 37 QSO dominant emission lines.

We have summarized the comparative numbers of O, W, and Q stars, novae and planetary nebulae to provide hints on their relative classification with respect to their evolution and the Hertzsprung-Russell diagram. The Hertzsprung-Russell diagram is a powerful tool to analyze and represent stellar evolution and has been used to determine the role of mass loss in the evolution of stars. In particular, we have considered the very important HR diagram of Ramachandra *et al* [39] for the stars of types O and early-B of the SMC Wing as it covers the temperature range up to 200 000 K, shows the Helium Zero-Age Main Sequence (He-ZAMS) and also exceptionally includes the Wolf-Rayet (WR) stars. In addition, their determination that hydrogen-depleted O stars with initial mass below ~30 $M_\odot$ evolve off the ZAMS to the right into colder red giants and supergiants, while hydrogen-depleted O stars with initial mass above $\sim$30 $M_\odot$ appear to always stay hot and veer off the ZAMS to the left to become Wolf-Rayet stars, beyond the O stars, is indicative of laser action in their stellar atmosphere.

We have reviewed the nucleosynthesis process that leads from massive O stars to WN and then WC Wolf-Rayet stars as a result of mass loss. We then considered the nucleosynthesis of oxygen in massive stars and found that ^{16}O oxygen has a significant presence in massive stars beyond the WC stage, until the generation of ^{28}Si, where it disappears. This has lead us to postulate more than one population of stars of type Q. One group identified by Varshni [27] with O VI and He II emission lines in their spectra implying much higher temperatures and positioning those QSO stars above the WR region in the HRD of Ramachandra *et al* [39].

The other group with emission lines dominated by temperatures in the O II and O III range, indicating a lower temperature range of QSO stars with a significant number of ionized oxygen lines, in addition to the nitrogen WN and carbon WC lines. We postulate that these QSO spectra, with a significant number of ionized O emission lines and some Si emission lines, correspond to unrecognized Wolf-Rayet stars, in particular WO stars and WSi pre-supernova stars, extending into lower temperatures. They in effect fill up the HRD of Ramachandra *et al* [39] in the range 16 000 K $< T_e <$ 73 000 K for stellar masses above $\sim$30 $M_\odot$. In that scenario, Q stars would be the end-state of the Wolf-Rayet evolution process and would account for their number larger than WN and WC W stars, prior to moving to the supernova state.

Received on February 15, 2021

References

1. Millette P. A. Laser Action in the Stellar Atmospheres of Wolf-Rayet Stars and Quasi-Stellar Objects (QSOs). *Prog. Phys.*, 2021, v. 17 (1), 71–82.
2. Menzel D. H. Laser Action in non-LTE Atmospheres. International Astronomical Union Colloquium, Volume 2: Spectrum Formation in Stars With Steady-State Extended Atmospheres, 1970, 134–137.
3. Gudzenko L. L., Shelepin L. A., Yakovlenko S. I. Amplification in recombining plasmas (plasma lasers). *Usp. Fiz. Nauk*, 1974, v. 114, 457. *Sov. Phys. - Usp.*, 1975, v. 17, 848.
4. Abbott D. C. and Conti P. S. Wolf-Rayet Stars. *Ann. Rev. Astron. Astrophys.*, 1987, v. 25, 113–150.
5. Crowther P. A. Physical Properties of Wolf-Rayet Stars. *Annu. Rev. Astron. Astrophys.*, 2007, v. 45, 177–219. arXiv: astro-ph/0610356v2.
6. Conti P. S. and Garmany C. D., De Loore C. and Vanbeveren D. The Evolution of Massive Stars: The Numbers and Distribution of O Stars and Wolf-Rayet Stars. *Astrophys. J.*, 1983, v. 274, 302–312.
7. Humphreys R. M. and Nichols M., Massey P. On the Initial Masses and Evolutionary Origins of Wolf-Rayet Stars. *Astron. J.*, 1985, v. 90, 101–108.
8. Conti P. S. and Vacca W. D. The Distribution of Massive Stars in the Galaxy: I. Wolf-Rayet Stars. *Astron. J.*, 1990, v. 100, 431–444.
9. Langer N., Hamann W.-R., Lennon M., Najarro F., Pauldrach A. Ẇ. A. and Puls J. Towards an understanding of very massive stars. A new evolutionary scenario relating O stars, LBVs [Luminous Blue Variables] and Wolf-Rayet stars. *Astron. Astrophys.*, 1994, v. 290, 819–833.
10. Crowther P. A., Smith L. J., Hillier D. J. and Schmutz W. Fundamental parameters of Wolf-Rayet stars. III. The evolutionary status of WNL [WN7 to WN9] stars. *Astron. Astrophys.*, 1995, v. 293, 427–445.
11. Hainich R., Rühling U., Todt H., Oskinova L. M., Liermann A., Gräfener G., Foellmi C., Schnurr O. and Hamann W.-R. The Wolf-Rayet stars in the Large Magellanic Cloud. A comprehendive analysis of the WN class. *Astronomy and Astrophysics*, 2014, v. 565, A27 (62 pp.).
12. Torres A. V. and Conti P. S. Spctroscopic Studies of Wolf-Rayet Stars. III. The WC Subclass. *Astrophys. J.*, 1986, v. 300, 379–395.
13. Koesterke L. and Hamann W.-R. Spectral analyses of 25 Galactic Wolf-Rayet stars of the carbon sequence. *Astron. Astrophys.*, 1995, v. 299, 503–519.
14. Barlow M. J. and Hummer D. J. The WO Wolf-Rayet stars. *Symposium – International Astronomical Union*, 1982, v. 99, 387–392.
15. Kingsburgh R. L., Barlow M. J. and Storey P. J. Properties of the WO Wolf-Rayet stars. *Astron. Astrophys.*, 1994, v. 295, 75–100.
16. van der Hucht K. A. The VIIth catalogue of galactic Wolf-Rayet stars. *New Astronomy Reviews*, 2001, v. 45, 135–232.
17. van der Hucht K. A. New Galactic Wolf-Rayet stars, and candidates (Research Note), An Annex to The VIIth Catalogue of Galactic Wolf-Rayet Stars. *Astronomy and Astrophysics*, 2006, v. 458, 453–459.
18. Galactic Wolf Rayet Catalogue. V1.25, www.pacrowther.staff.shef.ac.uk/WRcat/, Aug. 2020.
19. Mihalas D. Stellar Atmospheres, 2$^{\text{nd}}$ ed. W. H. Freeman and Co., San Francisco, 1978.
20. Mihalas D. and Weibel-Mihalas B. Foundations of Radiation Hydrodynamics, corr. ed. Dover Publications, New York, 1999, pp. 627–645.
21. Varshni Y. P., Talbot J., Ma Z. Peaks in Emission Lines in the Spectra of Quasars. In: Lerner E. J. and Almeida J. B., eds. 1$^{\text{st}}$ Crisis in Cosmology Conference, CCC-I. American Institute of Physics Conference Proceedings, v. 822, 2006.
22. Hewitt A., Burbidge G. A Revised and Updated Catalog of Quasi-stellar Objects. *Astrophys. J. Supp.*, 1993, v. 87, 451.
23. Willett C. S. Laser Lines in Atomic Species. In: Progress in Quantum Electronics, Vol. 1, Part 5, Pergamon, 1971.
24. Bennett W. R. Jr. Atomic Gas Laser Transition Data: A Critical Evaluation. IFI/Plenum, New York, NY, 1979.
25. Kramida A., Ralchenko Y., Reader J., and NIST ASD Team. NIST Atomic Spectra Database (ver. 5.8). National Institute of Standards and Technology, Gaithersburg, MD, 2020. //physics.nist.gov/asd, accessed 14 January 2021.

26. House L. L. Ionization Equilibrium of the Elements frpm H ro Fe. *Astrophys. J. Suppl.*, 1964, v. 8, 307–328.
27. Varshni Y. P. O VI and He II Emission Lines in the Spectra of Quasars. *Astrophys. Space Sci.*, 1977, v. 46, 443.
28. Maíz-Apellániz J. and Walborn N. R. A Galactic O Star Catalog. *Astrophys. J. Suppl.*, 2004, v. 151, 103–148.
29. Sota A., Maíz-Apellániz J., Walborn N. R., and Shida R. Y. The Galactic O Star Catalog V.2.0. arXiv: astro-ph/0703005.
30. Sota A., Maíz-Apellániz J., Walborn N. R., Alfaro E. J., Barrá R. H., Morrell N. I., Gamen R. C. and Arias J. L. The Galactic O-Star Spectroscopic Survey. I. Classification System and Bright Northern Stars in the Blue-Violet at R~2500. *Astrophys. J. Suppl.*, 2011, v. 193, 24–73.
31. Maíz Apellániz J., Pellerin A., Barbá R, H., Simón-Diaz S., Alfaro E. J., Morrell N. I., Sota A., Penatés Ordaz M. and Gallego Calvente A. T.. The Galactic O-Star Spectroscopic (GOSSS) and Northern Massice Dim Stars (NoMaDS) Surveys, the Galactic O-Star Catalog (GOSC), and Marxist Ghost Buster (MGB). arXiv: astro-ph/1109.1492.
32. Sota A., Maíz-Apellániz J., Morrell N. I., Barrá R. H., Walborn N. R., Gamen R. C., Arias J. L., and Alfaro E. J. The Galactic O-Star Spectroscopic Survey (GOSSS). II. Bright Southern Stars. *Astrophys. J. Suppl.*, 2014, v. 211, 10–93.
33. Maíz-Apellániz J., Moragón A. Á., de Zárate Alcarazo L. O. and the GOSSS team. The Galactic O-Star Catalog (GOSC) and the Galactic O-Star Spectroscopic Survey (GOSSS): current status. arXiv: astro-ph/1610.03320.
34. Parker Q. A., Acker A., Frew D. J., Hartley M., Peyaud A. E. J., Ochsenbein F., Phillipps S., Russell D., Beaulieu S. F., Cohen M., Köppen J., Miszalski B., Morgan D. H., Morris R. A. H., Pierce M. J. and Vaughan A. E. The Macquarie/AAO/Strasbourg Hα Planetary Nebula Catalogue: MASH. *Mon. Not. R. Astron. Soc.*, 2006, v. 373, 79–94.
35. Miszalski B., Parker Q. A., Acker A., Birkby J. L., Frew D. J. and Kovacevic A. MASH II: more planetary nebulae from the AAO/UKST Hα survey. *Mon. Not. R. Astron. Soc.*, 2008, v. 374, 525–534.
36. The Sloan Digital Sky Survey Quasar Catalog: sixteenth data release (DR16Q). www.sdss.org/dr16/algorithms/qso_catalog/, Aug. 2018.
37. scienceworld.wolfram.com/astronomy. accessed 24 Jan 2021.
38. Downes R. A., Webbink R. F., Shara M. M., Ritter H., Kolb U. and Duerbeck H. W. A Catalog and Atlas of Cataclysmic Variables: The Final Edition. *The Journal of Astronomical Data*, 2005, v. 11, 2.
39. Ramachandran V., Hamann W.-R., Oskinova L.M., Gallagher J. S., Hainich R., Shenar T., Sander A. A. C., Todt H. and Fulmer L. Testing massive star evolution, star formation history, and feedback at low metallicity. Spectroscopic analysis of OB stars in the SMC Wing. *Astronomy and Astrophysics*, 2019, v. 625, A104 (20 pp.). arXiv: astro-ph/1903.01762v2.
40. Hainich R., Pasemann D., Todt H., Shenar T., Sander A. and Hamann W.-R. Wolf-Rayet stars in the Small Magellanic Cloud. I. Analysis of the single WN stars. *Astronomy and Astrophysics*, 2015, v. 581, A21 (30 pp.).
41. Massey P. A Census of Massive Stars Across the Hertzsprung-Russell Diagram of Nearby Galaxies: What We Know and What We Don't. arXiv: astro-ph/0903.0155v2.
42. Meyer B. S., Nittler L. R., Nguyen A. N. and Messenger S. Nucleosynthesis and Chemical Evolution of Oxygen. *Review in Mineralogy & Geochemistry*, 2008, v. 68, 31–53. semanticscholar.org/paper/Nucleosynthesis-and-Chemical-Evolution-of-Oxygen-Meyer-Nittler.
43. Wallerstein G., Iben I. Jr., Parker P., Boesgaard A. M., Hale G. M., Champagne A. E., Barnes C. A., Käppeler F., Smith V. V., Hoffman R. D., Timmes F. X., Sneden C., Boyd R. N., Meyer B. S. and Lambert D. L. Synthesis of the elements in stars: forty years of progress. *Reviews of Moden Physics*, 1997, v. 69 (4), 995–1084.
44. Lengyel B. A. Introduction to Laser Physics. John Wiley, New York, 1966.
45. Willett C. S. Gas Lasers: Population Inversion Mechanisms. Pergamon Press, New York, 1974.
46. Millette P. A. Laser Action in C IV, N V, and O VI Plasmas Cooled by Adiabatic Expansion. University of Ottawa, Ottawa, ON, 1980. LaTeX typeset: researchgate.net/publication/283014713, 2015.
47. Varshni Y. P. Laser Action in Quasi-Stellar Objects? *Bull. Amer. Phys. Soc.*, 1973, v. 18, 1384.
48. Varshni Y. P. No Redshift in Quasi-Stellar Objects. *Bull. Amer. Astron. Soc.*, 1974, v. 6, 213.
49. Varshni Y. P. The Redshift Hypothesis and the Plasma Laser Star Model for Quasi-Stellar Objects. *Bull. Amer. Astron. Soc.*, 1974, v. 6, 308.
50. Varshni Y. P. Alternative Explanation for the Spectral Lines Observed in Quasars. *Astrophys. Space Sci.*, 1975, v. 37, L1.
51. Varshni Y. P. Electron Density in the Emission-Line Region of Wolf-Rayet Stars. *Astrophys. Space Sci.*, 1978, v. 56, 385.
52. Varshni Y. P. The Physics of Quasars. *Phys. Canada*, 1979, v. 35, 11.
53. Luyten W. J. A Search for Faint Blue Stars. Paper 50, University of Minnesota Observatory, Minneapolis, 1969.
54. Varshni Y. P. Proper Motions and Distances of Quasars. *Speculations in Science and Technology*, 1982, v. 5 (5), 521–532.
55. Mortlock D. J., Webster R. L. and Francis P. J. Binary Quasars. *Mon. Not. R. Astron. Soc.*, 1999, v. 309, 836–846.
56. Hennawi J. F., et al. Binary Quasars in the Sloan Digital Sky Survey: Evidence for Excess Clustering on Small Scales. *Astronomical Journal*, 2006, v. 131, 1–23.

Physics of Numbers as Model of Telepathic Entanglement

Hartmut Müller, Renata Angeli, Roberta Baccara, Rose Line Hofmann,
Simona Muratori, Olga Nastasi, Giuliana Papa, Francesca Santoni,
Claudio Venegoni, Francesco Zanellati, Leili Khosravi

E-mail: hm@interscalar.com

The physics of transcendental numbers leads to a fractal scalar field that causes numeric entanglements affecting any type of interaction. In this paper, we apply this our approach to the analysis of telepathic communication in both aspects, the theoretical and experimental.

Introduction

The history of science is replete with confident proclamations about all sorts of impossible things like flying machines heavier than air, and most of those proclamations have proven to be hilariously or poignantly wrong. So the current paradigm declares also telepathy to be impossible [1].

The term 'telepathy' comes from the Greek 'tele' meaning 'distant' and 'pathos' meaning 'feeling, perception, experience' and can be defined [2] as the transmission of information from one person to another without using any known human sensory channel or physical interaction.

Introduced by the British scholar Frederic W. H. Myers in 1882, 'telepathy' substituted the earlier term 'thought transference' in psychology. The concept of telepathy was originally more an attempt to objectify and detach the concept of thought transference from its connection with spiritism, media and belief in ghosts.

Telepathy challenges the scientific understanding of experience, that David Chalmers [3] has termed the 'hard problem' of consciousness. Indeed, centuries of philosophical disputes did not explain the nature of consciousness. Aside from recognizing that consciousness differs from matter in many ways, there is no scientific consensus.

However, the dominant view in recent time is more materialistic than ever before: consciousness is thought to emerge from highly complex biological processes, which in turn are based ultimately on interactions between subatomic particles.

Roger Penrose and Stuart Hameroff [4] hypothesize that consciousness originates from quantum processing in neuron dendritic spine microtubules.

Shan Gao [5] analyzes the role of consciousness during quantum measurement process and supposes quantum nonlocality as model of telepathic communication. Huping Hu and Maoxin Wu [6, 7] hypothesize that consciousness is intrinsically connected to quantum spin in the sense that nuclear and electron spin is the 'mind-pixel' and the unity of mind is presumably achieved by entanglement of these mind-pixels. They assume [8] that spin is the primordial process in non-spatial and non-temporal pre-spacetime being the manifestation of quantum entanglement, implying instantaneous interconnectedness of all matters in the universe through gravity and consciousness. As well, George Williams [9] supposes the existence of a non-local proto-conscious field that underlies both matter and consciousness. Within the Global Consciousness Project of the Princeton Engineering Anomalies Research Laboratory at the Princeton University, the Rodger Nelson group [10] demonstrated that human consciousness interacts with physical random event generators [11], causing them to produce nonrandom patterns associated with special states of group consciousness.

In our research we focus on the physics of numbers as approach to study the physical consequences of arithmetic properties of numbers being ratios of measured quantities. In [12] we have shown that this approach leads to a fractal scalar field that causes numeric entanglements affecting any type of interaction including gravitation [13]. In this paper, we apply our approach to the analysis of telepathic communication in both aspects, the theoretical and experimental.

Theoretical Approach

Measurement is the source of scientific data that allow for developing and proofing theoretical models of the reality. The result of a measurement is the ratio of two quantities where one of them is the reference quantity called unit of measurement. All that can be measured – space, time, energy, mass – is quantity. Numbers are symbols of quantity. Despite their non-materiality, numbers represent a reality that has unlimited power and produces physical effects. These effects are a subject of study in the physics of numbers.

On the one hand, numbers appear as created by intellect, on the other, our intellect cannot manipulate them, for example, avoid the appearance of primes when counting, or design a cube and a sphere both of the same volume. Indeed, measuring, counting and calculating are inherent abilities of all things. Even atoms have to configure the number of electrons on each energy level. Thus, the universality of the numbers suggests that they are not anthropogenic, but cosmogenic.

Distances, durations, angles, velocities – when measured, first they are real numbers, and only when applied to models they can become vectors. Real numbers are scalars, and scaling is the process that creates them. Indeed, when we observe something from a far scale, we cannot distinguish details. Different objects appear as identical and we cannot anymore individuate them. The abundance of properties of the objects reduces to their number that follows the laws of arithmetic or the laws of statistics.

Extreme scaling is the process that creates numbers and can possibly even release objects from their materiality. The scale of electrons is in the range of picometer. Protons and electrons appear to be elementary just because the difference between the observer's macroscale and the subatomic scale is huge. This is why they behave like numbers and their properties appear quantized following the laws of quantum statistics.

Numbers are omnipresent and therefore, non-local. This non-locality of the numbers might be the true cause of the quantum physical entanglement that Albert Einstein called 'spooky action at a distance'. In this context, all electrons and protons are identical because there is probably only one electron and only one proton that can materialize everywhere. In the same meaning there is only one number e=2.71828... and only one number π=3.14159... that can materialize anytime and anywhere.

Max Planck's discovery that the energy $E=\hbar\omega$ of a photon depends only on a *number* that is its frequency ω, is a key event in the history of physics. From this discovery, quantum physics was born. As the energy of a quantum oscillator increases with its frequency, every additional increase of the frequency requires more and more energy. Probably, this process leads to the emergence of a resistance that appears as inertia. Indeed, the frequency $7.8 \cdot 10^{20}$ Hz defines the threshold where electrons can form. Surpassing the threshold of $1.4 \cdot 10^{24}$ Hz, protons can form. In [14] we introduced scaling as mechanism of particle mass generation, alternative to the Higgs model. In [15] we have demonstrated that it is the transcendence of Euler's number that stabilizes the thresholds of materialization including the proton-to-electron ratio.

In the framework of the physics of numbers, all structures and processes in the universe are materializations of numeric relationships. Within this our approach, we significantly extend the meaning of quantum entanglement in the sense of an instantaneous connectivity that originates from the divisibility of numbers. The meaning of this connectivity is that, for example, the n^{th} cycle of a given process has something in common with the n^{th} cycle of any other process, independently of its nature, duration or location.

This kind of 'numeric entanglement' is a consequence of the divisibility of the number n being the index of the n^{th} cycle of a periodical process. It has nothing to do with resonance or simultaneity, but with scaling; it is a connectivity that does not depend on temporal coincidences or spatial distances.

Let us imagine two periodic processes, one occurs on Earth and another occurs on Kepler 452b that is 1400 light years away in the Cygnus constellation of the Milky Way. Because of the huge spatial and temporal distance, they cannot be synchronized by the speed of light. By the way, that's exactly why probably nobody in the Galaxy uses radio signals or other forms of light for interstellar communication. Nevertheless, both periodic processes are numerically connected, and this circumstance allows for communication.

In [12] we have demonstrated that the physics of transcendental numbers leads to a fractal scalar field that affects any type of physical interaction including gravitation. In this paper, we hypothesize that this field causes numeric entanglements making possible connectivity associated with telepathy or other forms of extrasensory perception. But first, now we are going to derive this fractal scalar field from the physics of transcendental numbers.

In physics of numbers [16], the difference between rational, irrational algebraic and transcendental numbers is not only a mathematical task, but it is also an essential aspect of stability in complex dynamic systems. While integer frequency ratios provide parametric resonance interaction that can destabilize a system [17, 18], it is transcendental numbers that define the preferred ratios of quantities which avoid destabilizing resonance interaction [15]. In this way, transcendental ratios of quantities can sustain the stability of periodic processes in complex dynamic systems.

Among all transcendental numbers, Euler's number e = 2.71828... is unique, because its real power function e^x coincides with its own derivatives. In the consequence, Euler's number allows inhibiting resonance interaction regarding any interacting periodic processes and their derivatives.

Alexandr Khinchin [19] demonstrated that any real number has a biunique representation as a continued fraction. Applying this to the real argument x of the natural exponential function e^x, we get:

$$x = \langle n_0; n_1, n_2, \ldots, n_k \rangle. \qquad (1)$$

We use angle brackets for continued fractions. All denominators $n_1, n_2, \ldots, n_k$ including the free link n_0 are integer. The numerators equal 1. The length of the continued fraction is given by the number k of layers.

The canonical form (all numerators equal 1) does not limit our conclusions, because every continued fraction with partial numerators different from 1 can be transformed into a canonical continued fraction using the Euler equivalent transformation [20]. With the help of the Lagrange [21] transformation, every continued fraction with integer denominators can be represented as a continued fraction with natural denominators that is always convergent [22].

Naturally, the rational eigenvalues of the finite continued fractions (1) have a fractal distribution. The first layer is given by the truncated after n_1 continued fraction:

$$x = \langle n_0; n_1 \rangle = n_0 + \frac{1}{n_1}.$$

The denominator n_1 follows the sequence of integer numbers ±1, ±2, ±3 etc. The second layer is given by the truncated after n_2 continued fraction:

$$x = \langle n_0; n_1, n_2 \rangle = n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2}}.$$

Figure 1 shows the first and the second layer in comparison. As we can see, reciprocal integers $\pm 1/2, \pm 1/3, \pm 1/4, \ldots$ are the attractor points of the distribution. In these attractors, the distribution density always reaches a local maximum. Integers $0, \pm 1, \ldots$ are the main attractors of the distribution.

Now let's remember that we are observing the fractal distribution of rational values $x = \langle n_0; n_1, n_2, \ldots, n_k \rangle$ of the real argument x of the natural exponential function e^x. What we see is the fractal distribution of transcendental numbers of the type $\exp(\langle n_0; n_1, n_2, \ldots, n_k \rangle)$ on the natural logarithmic scale. Near integer exponents, the distribution density of these transcendental numbers is maximum. Consequently, for integer and rational exponents x, the natural exponential function e^x defines attractor points of transcendental numbers and create islands of stability.

Figure 1 shows that these islands are not points, but ranges of stability. Integer exponents $0, \pm 1, \pm 2, \pm 3, \ldots$ are attractors which form the widest ranges of stability. Half exponents $\pm 1/2$ form smaller islands, one third exponents $\pm 1/3$ form the next smaller islands and one fourth exponents $\pm 1/4$ form even smaller islands of stability etc.

For rational exponents, the natural exponential function is always transcendental [23]. Increasing the length k of the continued fraction (1), the density of the distribution of transcendental numbers of the type $\exp(\langle n_0; n_1, n_2, \ldots, n_k \rangle)$ is increasing as well. Nevertheless, their distribution is not homogeneous, but fractal. Applying continued fractions and truncating them, we can represent the real exponents x of the natural exponential function e^x as rational numbers and make visible their fractal distribution.

The application of continued fractions doesn't limit the universality of our conclusions, because continued fractions deliver biunique representations of all real numbers including transcendental. Therefore, the fractal distribution of transcendental eigenvalues of the natural exponential function e^x of the real argument x, represented as continued fraction, is an inherent characteristic of the number continuum. This characteristic we call the *Fundamental Fractal* [24].

In physical applications, the natural exponential function e^x of the real argument x is the ratio of two physical quantities where one of them is the reference quantity called unit of measurement. Therefore, we can rewrite the equation (1):

$$\ln(\mathrm{X}/\mathrm{Y}) = \langle n_0; n_1, n_2, \ldots, n_k \rangle \tag{2}$$

where X is the measured physical quantity and Y the unit of measurement. In this way, the natural exponential function e^x of the rational argument $x = \langle n_0; n_1, n_2, \ldots, n_k \rangle$ generates the set of preferred ratios X/Y of quantities which avoid destabilizing resonance and provide the lasting stability of real systems regardless of their complexity.

Therefore, we expect that periodic processes in real systems prefer frequency ratios close to Euler's number and its rational powers. Consequently, the logarithms of their frequency ratios should be close to integer $0, \pm 1, \pm 2, \ldots$ or rational values $\pm 1/2, \pm 1/3, \pm 1/4, \ldots$

In [12] we verified the model claims on the gravitational constants and the periods of orbital and rotational motion of the planets, planetoids and large moons of the solar system as well as the orbital periods of exoplanets and the gravitational constants of their stars.

Naturally, the Fundamental Fractal (2) of transcendental stability attractors does not materialize in the scale of planetary systems only. At subatomic scale, it stabilizes the proton-to-electron ratio and in this way, allows the formation of stable atoms and complex matter.

Scale relations in particle physics [14] obey the same Fundamental Fractal (2), without any additional or particular settings. The proton-to-electron frequency ratio approximates the Fundamental Fractal at the first layer that could explain their exceptional stability [25]:

$$\ln\left(\frac{\omega_p}{\omega_e}\right) = \ln\left(\frac{1.42549 \cdot 10^{24}\,\text{Hz}}{7.76344 \cdot 10^{20}\,\text{Hz}}\right) \simeq 7 + \frac{1}{2} = \langle 7; 2 \rangle.$$

ω_p and ω_e are the proton and electron angular frequencies. In the consequence of the ratio $\exp(7 + 1/2)$, the scaling factor $\sqrt{e} = 1.64872\ldots$ connects attractors of proton stability with similar attractors of electron stability in alternating sequence. Figure 2 demonstrates this situation on the first layer of the Fundamental Fractal (1), and one can see clearly that among the integer or half, only the attractors $\pm 1/3$, $\pm 1/4$ and $\pm 1/6$ are common. In these attractors, proton stability is supported by electron stability and vice versa, so we expect that they are preferred in real systems. As we have shown in our paper [12], planetary systems make extensive use of these common attractors.

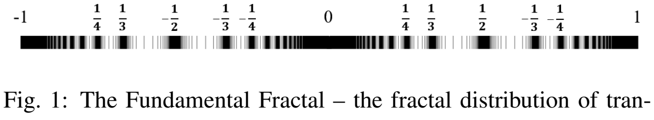


Fig. 1: The Fundamental Fractal – the fractal distribution of transcendental numbers of the type e^x with $x = \langle n_0; n_1, n_2, \ldots, n_k \rangle$ on the natural logarithmic scale for $k = 1$ (first layer above) and for $k = 2$ (second layer below) in the range $-1 \leqslant x \leqslant 1$.

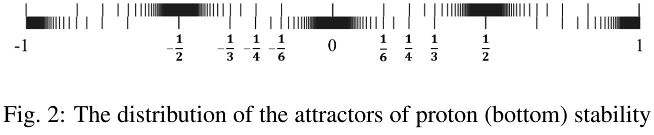


Fig. 2: The distribution of the attractors of proton (bottom) stability in the range $-1 < x < 1$ of the attractors of electron (top) stability. Natural logarithmic representation.

Property	Electron	Proton
$E = mc^2$	0.5109989461(31) MeV	938.2720813(58) MeV
$\omega = E/\hbar$	$7.76344 \cdot 10^{20}$ Hz	$1.42549 \cdot 10^{24}$ Hz
$\tau = 1/\omega$	$1.28809 \cdot 10^{-21}$ s	$7.01515 \cdot 10^{-25}$ s
$\lambda = c/\omega$	$3.86159 \cdot 10^{-13}$ m	$2.10309 \cdot 10^{-16}$ m

Table 1: The basic set of the physical properties of the electron and proton. Data from Particle Data Group [29]. Frequencies, oscillation periods and wavelengths are calculated.

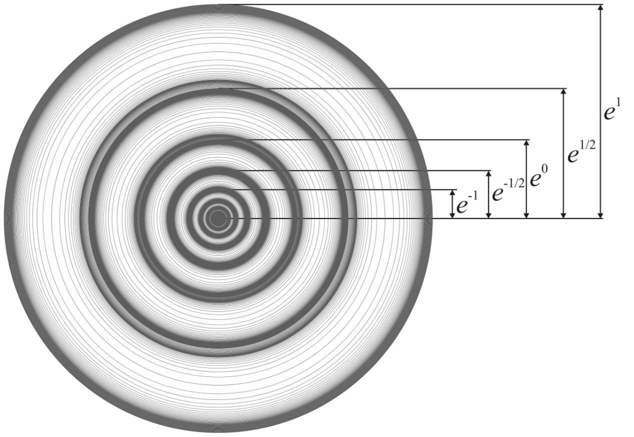


Fig. 3: The Fundamental Field with equipotential surfaces of both proton and electron attractors of stability in the linear 2D-projection for $k = 1$ in the range $-1 < x < 1$.

The spatio-temporal projection of the Fundamental Fractal (2) is a fractal scalar field of transcendental attractors, the *Fundamental Field* [26]. The connection between the spatial and temporal projections of the Fundamental Fractal is given by the speed of light c = 299792458 m/s. The constancy of c makes both projections isomorphic, so that there is no arithmetic or geometric difference. Only the units of measurement are different. In [27] we have shown that the constancy of the speed of light is a consequence of the stabilizing function of Euler's number.

The exceptional stability of the electron and proton predestinate their physical characteristics as fundamental units. Table 1 shows the basic set of electron and proton units that we consider as a fundamental metrology (c is the speed of light in a vacuum, $\hbar$ is the Planck constant). In [24] was shown that the fundamental metrology (tab. 1) is completely compatible with Planck units [28].

The Fundamental Field is topologically 3-dimensional, it is a fractal set of embedded spherical equipotential surfaces. Figure 3 shows the linear 2D-projection $\exp(1/n_1)$ of the first layer of the Fundamental Field with both proton and electron attractors of stability. Figure 2 shows the same interval in the logarithmic representation.

In [30] we have shown that the frequency boundaries of the brain activity ranges Delta, Theta, Alpha, Beta and Gamma do not appear as to be accidental, but correspond with attractors of proton and electron stability of the Fundamental Fractal (2). In this way, Euler's number determines temporal scales of stability of the central nervous system. Indeed, mammals including human have electrical brain activity [31] of the Theta type in the frequency range between 3 and 7 Hz, of Alpha type between 8 and 13 Hz and Beta type between 14 and 37 Hz. Below 3 Hz the brain activity is of the Delta type, and above 37 Hz the brain activity changes to Gamma. The frequencies 3.0 Hz, 8.2 Hz, 13.5 Hz and 36.7 Hz define the boundaries. The logarithms of their ratios are close to integer and half values:

$$\ln\left(\frac{8.2}{3.0}\right) = 1.00; \qquad \ln\left(\frac{13.5}{8.2}\right) = 0.50; \qquad \ln\left(\frac{36.7}{13.5}\right) = 1.00.$$

The correspondence of the boundary frequency ratios with integer and half powers of Euler's number evidences that the stability of the frequency boundaries is essential for brain activity. In fact, Theta-Alpha or Alpha-Beta violence can cause speech and comprehension difficulties, depression and anxiety disorders. Figure 6 shows how precisely the frequency boundaries of all subranges of brain activity correspond with main attractors of proton and electron stability.

Furthermore, similar boundary frequencies we find in the Earth's electromagnetic field spectrum, for example the fundamental Schumann mode 7.8 Hz. Solar X-ray bursts can cause variations of the Schumann resonances [32]. In this case, the fundamental increases up to 8.2 Hz reaching exactly the stable Theta-Alpha boundary. The second Schumann mode 13.5 Hz coincides precisely with the Alpha-Beta boundary. It is remarkable that solar activity affects this mode much less or does not affect it at all because of its Euler stability. The third Schumann mode 20.3 Hz must increase up to 22.2 Hz for reaching the next island of electron stability. By the way, such an increase is observed recently. Schumann resonances occur up to 60 Hz in order to reach the subsequent island of electron stability.

The coincidence of the boundary frequencies of brain activity with Schumann resonances demonstrates how precisely the electrical activity of biological systems is embedded in the electromagnetic activity of the Earth. Important to know that Euler's number and its roots make possible this embedding, because they are attractors of transcendental numbers and form islands of stability. They allow for exchanging information between systems of very different scales – the biophysical and the geophysical.

Here and in the following we use the letter E for attractors of electron stability, and the letter P for attractors of proton stability. For instance, the attractor $E\langle -48 \rangle$ dominates the Delta activity range while $E\langle -45 \rangle$ dominates the Beta activity

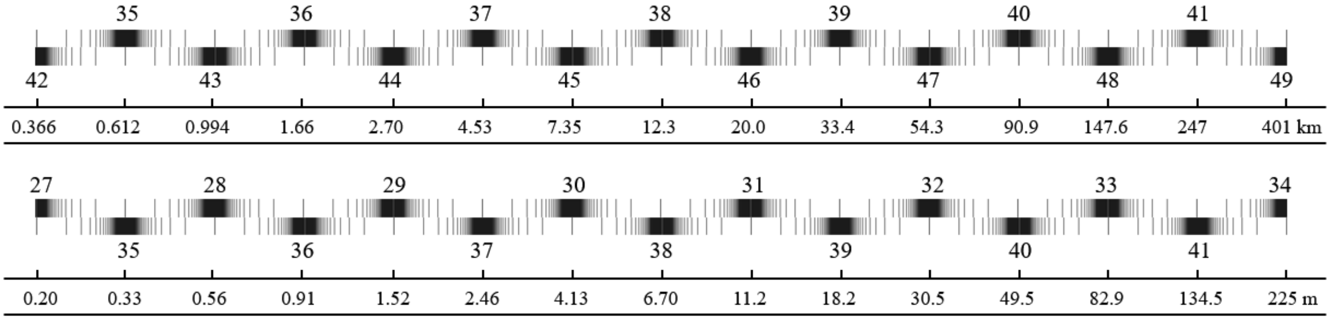


Fig. 4: Radii of equipotential surfaces of the Fundamental Field (Fig. 3) and the corresponding attractors of electron and proton stability in the natural logarithmic representation.

range. The Theta and Gamma activity ranges are dominated by the attractors P⟨−54⟩ and P⟨−51⟩ of proton stability. These logarithms are multiples of 3. Low Delta, High Delta, Alpha and Low Gamma are transition ranges, which boundaries are defined by both, attractors of electron and proton stability. For instance, dividing the Theta – Alpha boundary frequency 8.2 Hz by the electron angular frequency, we can see how precisely it matches with the attractor E⟨−46⟩ of electron stability:

$$\ln\left(\frac{8.2\,\mathrm{Hz}}{7.76344\cdot 10^{20}\,\mathrm{Hz}}\right) = -46.00.$$

The correspondence of the boundary frequencies with attractors of proton and electron stability evidences that *quantum* physical stability of the frequency boundaries is essential for brain activity. Perhaps, this could also indicate that brain-to-brain entanglements are possible. Probably, the attractor frequencies are the key. To verify this hypothesis, we designed an experimental setup that we describe in the following.

Experimental Setup

The experiments of telepathic communication described in this paper were performed continuously over a period of four years. The participants have decades of experience in meditation, and are married couples respective good friends. They took turns in their roles as sender and receiver. During the first year, a sender usually tried to transmit the information about an arbitrarily chosen object – an apple, stone, ring or painting – that the receiver had to identify and describe in written form and draw.

Fig. 5: The electric skin potential of the sender (black curve) and the receiver (grey curve) measured with transient recorders of two DSOs. The resolution is 100 measurements per second. The time-window of each graphic is one second.

For reduction of the interference of electrical brain activity by low frequency external electromagnetic fields, a part of the receivers and/or senders applied hypo-electromagnetic constructions made of 1/16 aluminum sheet, similar to the described in [13] polyhedrons, as helmets. Larger constructions of the same material were used to stay inside a hypo-electro-magnetic space where modulated red light was applied as well. For LED modulation, the frequencies 3, 5, 13, 23, 37, 61 or 101 Hz (fig. 6) of electron and proton stability were chosen. The dimensions of the structures coincide with the radii P⟨35⟩=0.33 m, E⟨28⟩=0.56 m, P⟨36⟩=0.91 m, E⟨29⟩=1.52 m and P⟨37⟩=2.46 m of equipotential surfaces of the Fundamental Field.

The distance between sender and receiver partly was chosen in accordance with radii of main equipotential surfaces of the Fundamental Field. Fig. 4 shows the complete spectrum of sizes and distances that was applied in the experiments.

The durations of the transmission setup stages were chosen in accordance with main temporal attractors of the Fundamental Fractal (fig. 7). In the first generation of the experiments, the long version of the transmission setup stages was chosen taking 15 minutes. Then, in the next generations of experiments, the short version that takes 5 minutes only was established.

The protocols of these experiments contain information not only about the very object, its origin, meaning and background, but also about the physical and mental state of the sender, colors of dress and other details of the environment, and of course, time and geographic location. Particular attention was paid to the perception of time. During the experiments, the participants usually were at home in Milan, Malnate, Ferrara, Ravenna, Arezzo, Spigno Saturnia, Castel di Fiori or Citta della Pieve, so that the telepathic communication did occur over large distances up to 420 km beeline.

In experiments over short distances up to 7 meters, the electric skin potential of the participants was measured. Two digital storage oscilloscopes were used. During the experiment, the participants were contacting the measurement electrodes of the DSO with a finger.

Figure 5 shows the signals of the sender (black curve) and the receiver (grey curve) measured with the transient recorder of the DSO. The resolution of the transient recorder is 100 measurements per second. The graphic at the top shows one second of the alignment during the second minute after the start of the experiment. The middle graphic shows one second of the initial phase of the entanglement during the third minute, and the graphic at the buttom shows one second of the entanglement during the fourth minute.

The unexpected success and the frequent cases of very detailed description of the objects and even the sender's environment inspired to continue the experiments under more controlled conditions.

Therefore, in the 2nd generation of experiments, the arbitrarily chosen object was substituted by a simple geometric form. The sender chooses one of four easily distinguishable forms – cross, triangle, square or circle – for transmission, and the receiver must identify it.

Furthermore, for controlling the dependence of transmission success on the number of participants, the experiments were carried out with two and more receivers. In the 3^{d} generation, the geometric forms were substituted by six domino number configurations (fig. 8).

In the 4th generation of experiments, the geometric forms were substituted by Chladni patterns. Fine sand particles accumulate in nodal patterns on the surface of vibrating metal plates, as described by Galileo Galilei (1630), Robert Hook (1680) and Ernst Chladni (1787). The emerging patterns depend only on the geometry of the plate and the vibration frequency of the particles, and do not depend on their mass or chemical composition. This characteristic remembers gravity – as the acceleration of free fall does not depend on the mass of the test body or its chemical composition.

For the experiments, Chladni patterns (fig. 9) emerging on square plates vibrating with the frequencies of 150, 175, 179, 400 and 525 Hz were used. On the Fundamental Fractal, these frequencies are distributed around the main nodes $E\langle -43\rangle$ and $E\langle -42\rangle$ of electron stability, as fig. 10 shows.

The 5th generation of experiments dealt with 5 kingdoms of nature – human, animal, vegetal, mineral and celestial bodies. The transmission time extended throughout the day without specific mental focus. The sender shall transmit the idea of a concrete representative of one of these 5 kingdoms that the receiver has to identify as detailed as possible. If the kingdom of the transmitted representative was identified correctly (for example, animal), the coefficient of success was counted as 1/5, and if the representative was identified (for example, lion), the transmission was double rated. In the 6th generation of experiments, the sender tried to transmit one of five 'states of soul'. The first set of such states included courage, patience, joy, beauty and kindness, and the second set included enthusiasm, calm, trust, gratitude and benevolence. The qualities have been modified to avoid falling into monotony due to the fact that after about a month the participants felt that the exact perceptions decrease.

Results

A total of 242 experiments were carried out from September 2016 to November 2020, and the unexpected high rate of success let the participants believe in the reality of telepathy. With growing up experience, the receiver felt to be capable observing the world through the eyes of the sender. Obviously, every kind of information can be transmitted and is not limited by emotions or feelings, but can include detailed descriptions of real objects as well as numbers, regular forms and even sophisticated patterns or paintings.

The chance probability that the receiver is able to correctly guess one of five possibilities is 1/5 = 20%. However, the combined hit rates in our 114 experiments of that type was 72%. Statistically, this excess would never occur by chance; it corresponds to odds against chance of billions to one. This fact indicates that sender and receiver had shared indeed the same information. Such a high rate of success is not typical for the branch. As reported in [1, 33], good hit rates typically exceed the statistical expectation by 3 – 12%. Therefore, a possible significance of special conditions is obvious. Friendship and love are powerful connectors, and our research would not be necessary for a confirmation. Although these factors of success were always present in our research, they alone cannot explain the exceptionally high hit rates.

Initially, the hit rates did correlate with the distance between sender and receiver depending on the vicinity to a main equipotential surface of the Fundamental Field, but with increasing experience, this factor did lose its significance. As well, hypo-electromagnetic conditions initially did support the occurrence of telepathic entanglements significantly. Also modulated light initially did it, if the modulation frequency did correspond with an attractor of electron or proton stability. Despite this development, the statistics of the experiments evidence the permanent significance of the temporal and spa-

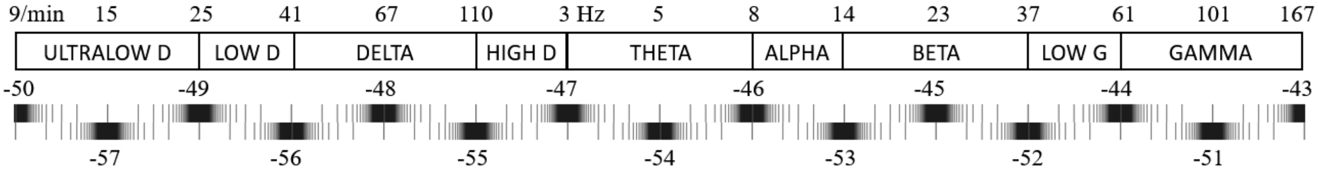


Fig. 6: The frequency boundaries of the brain activity ranges and the corresponding attractors of proton (below) and electron (above) stability of the Fundamental Fractal (2) in the natural logarithmic representation.

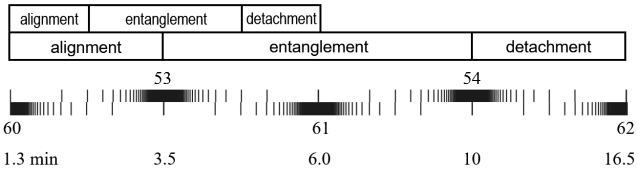


Fig. 7: The duration of transmission setup stages in minutes (below) and the corresponding attractors of proton and electron stability.

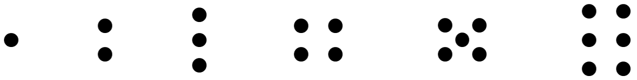


Fig. 8: Domino number configurations applied in the 3^{th} generation.

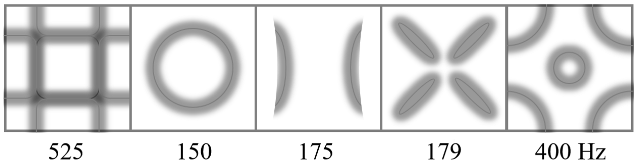


Fig. 9: Chladni patterns emerging on a vibrating square metal plate driven with the frequencies 150, 175, 179, 400 and 525 Hz.

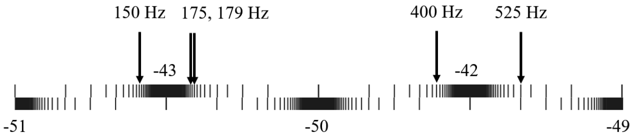


Fig. 10: The applied frequencies 150, 175, 179, 400 and 525 Hz and the corresponding attractors of proton (below) and electron (above) stability of the Fundamental Fractal (2) in the natural logarithmic representation.

tial attractors for the dynamics of the telepathic entanglement. Closer to the end of the four years' experience, indeed, the durations of the transmission setup stages automatically obeyed the Fundamental Fractal in a very natural way.

Starting with the 2^{nd} generation, the experiments were carried out with two and more receivers. This fact in particular enables a more precise model selection and clearly shows that telepathy is not limited by individual entanglement.

Conclusion

Finally, our experiments helped to discard some conventional hypotheses provided to explain telepathy. Considering the empirical fact that electromagnetic isolation supports telepathic entanglement, today we discard the idea that telepathy is based on electromagnetic waves. It would also be a joke to think that gravitational waves could be responsible for telepathy. We suppose that telepathy has nothing to do with signal transmission. In some cases indeed, the receiver got the information *before* the sender decided to share it.

We hypothesize that besides of electromagnetic and gravitational fields, there is another long-range phenomenon – the Fundamental Field – that is of pure numeric origin and nonmaterial, like consciousness. This Fundamental Field could turn out to be a primordial field from which consciousness originates. Being not limited by any physical process, the Fundamental Field causes numeric entanglements affecting any type of interaction.

Within our approach, telepathy is an access to a common quantum physical pool of information. Thanks to the nonlocality of this pool, every telepathist can get the required information. Any process, any event updates the quantum physical information pool automatically. No sender is needed. Accessing the pool, the participant A seeks for information that is related to the participant B.

Obviously, special conditions can facilitate this access. In our experiments, those conditions were always related to the Fundamental Field. Therefore, we propose numeric entanglement as model of telepathic sharing of information.

Acknowledgements

The authors are grateful to the Community of Living Ethics for permanent support on all stages of the study.

Submitted on February 27, 2021

References

1. Radin D. Thinking about telepathy. *Think*, March 2003, doi: 10.1017/S1477175600000415.
2. Telepathy. www.britannica.com
3. Chalmers D. Facing up to the problem of consciousness. *Journal of Consciousness Studies*, (1995).
4. Hameroff S., Penrose R. Conscious Events as Orchestrated Space-Time Selections. *NeuroQuantology*, 2003, v. 1, 10–35.

5. Gao Shan. A primary quantum model of telepathy. *SSRN Electronic Journal*, doi: 10.2139 / ssrn.3076089, (2002).
6. Hu H. P., Wu M. X. Spin-Mediated Consciousness Theory. *Med. Hypotheses*, 2004, v. 63, 633–646; arXiv: quant-ph/0208068.
7. Hu H. P., Wu M. X. Nonlocal Effects of Chemical Substances on the Brain Produced through Quantum Entanglement. *Progress in Physics*, 2006, v. 3, 20–26.
8. Hu H. P., Wu M. X. Thinking outside the box II: The Origin, Implications and Applications of Gravity and its role in Consciousness. *Neuro-Quantology*, 2006, v. 4, 5–16; Cogprints: ID4581.
9. Williams G. Psi and the Problem of Consciousness. *The Journal of Mind and Behavior*, 2013, v. 34, no 3, 4, 259–284.
10. Nelson R. D. Correlation of Global Events with REG Data. *The Journal of Parapsychology*, 2001, v. 65, 247–271.
11. Nelson R. D., Bradish G. J., Dobyns Y. H. Random Event Generator Qualification, Calibration, and Analysis. Princeton University, PEAR 89001, (1989).
12. Müller H. Physics of Transcendental Numbers Meets Gravitation. *Progress in Physics*, 2021, v. 17, 83–92.
13. Müller On the Acceleration of Free Fall inside Polyhedral Structures. *Progress in Physics*, 2018, v. 14, 220–225.
14. Müller H. Emergence of Particle Masses in Fractal Scaling Models of Matter. *Progress in Physics*, 2012, v. 8, 44–47.
15. Müller H. The Physics of Transcendental Numbers. *Progress in Physics*, 2019, v. 15, 148–155.
16. Müller H. On the Cosmological Significance of Euler's Number. *Progress in Physics*, 2019, v. 15, 17–21.
17. Dombrowski K. Rational Numbers Distribution and Resonance. *Progress in Physics*, 2005, v. 1, no. 1, 65–67.
18. Panchelyuga V.A., Panchelyuga M. S. Resonance and Fractals on the Real Numbers Set. *Progress in Physics*, 2012, v. 8, no. 4, 48–53.
19. Khintchine A.Ya. Continued fractions. University of Chicago Press, Chicago, (1964).
20. Skorobogatko V. Ya. The Theory of Branched Continued Fractions and mathematical Applications. Moscow, Nauka, (1983).
21. Lagrange J. L. Additions aux elements d'algebre d'Euler. (1798).
22. Markov A. A. Selected work on the continued fraction theory and theory of functions which are minimum divergent from zero. Moscow–Leningrad, (1948).
23. Hilbert D. Über die Transcendenz der Zahlen e und π. *Mathematische Annalen*, 1893, v. 43, 216–219.
24. Müller H. Scale-Invariant Models of Natural Oscillations in Chain Systems and their Cosmological Significance. *Progress in Physics*, 2017, v. 13, 187–197.
25. Müller H. Global Scaling. The Fundamentals of Interscalar Cosmology. *New Heritage Publishers*, Brooklyn, New York, USA, ISBN 978-0-9981894-0-6, (2018).
26. Müller H. Quantum Gravity Aspects of Global Scaling and the Seismic Profile of the Earth. *Progress in Physics*, 2018, v. 14, 41–45.
27. Müller H. The Cosmological Significance of Superluminality. *Progress in Physics*, 2019, v. 15, 26–30.
28. Astrophysical constants. Particle Data Group, www.pdg.lbl.gov
29. Tanabashi M. et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018), www.pdg.lbl.gov
30. Müller H. Chain Systems of Harmonic Quantum Oscillators as a Fractal Model of Matter and Global Scaling in Biophysics. *Progress in Physics*, 2017, v. 13, 231–233.
31. Tesche C. D., Karhu J. Theta oscillations index human hippocampal activation during a working memory task. *PNAS*, 2000, v. 97, no. 2, 919–924.
32. Roldugin V. C. et al. Schumann resonance frequency increase during solar X-ray bursts. *Journal of Geophysical Research*, 2014, v. 109, A01216.
33. De Peyer J. Telepathic Entanglements: Where are we Today? Psychoanalytic Dialogues, 2014, v. 24, 109–121.

Dirac 4×1 Wavefunction Recast into a 4×4 Type Wavefunction

G. G. Nyambuya[1,2]

[1]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.
[2]The Copperbelt University, School of Mathematics and Natural Sciences – Department of Physics,
Fundamental Theoretical and Astrophysics Group, P. O. Box 21692, Jambo Drive – Riverside, Kitwe, Republic of Zambia.
E-mail: physicist.ggn@gmail.com

As currently understood, the Dirac theory employs a 4×1 type wavefunction. This 4×1 Dirac wavefunction is acted upon by a 4×4 Dirac Hamiltonian operator, in which process, four independent particle solutions result. Insofar as the real physical meaning and distinction of these four solutions, it is not clear what these solutions really mean. We demonstrate herein that these four independent particle solutions can be brought together under a single roof wherein the Dirac wavefunction takes a new form as a 4×4 wavefunction. In this new formation of the Dirac wavefunction, these four particle solutions precipitate into three distinct and mutuality dependent particles (ψ_L, ψ_N, ψ_R) that are eternally bound in the same region of space. Given that quarks are readily found in a mysterious threesome cohabitation-state eternally bound inside the proton and neutron, we make the suggestion that these Dirac particles (ψ_L, ψ_N, ψ_R) might be quarks. For the avoidance of speculation, we do not herein explore this idea further but merely present it as a very interesting idea worthy of further investigation. We however must say that, in the meantime, we are looking further into this very interesting idea, with the hope of making inroads in the immediate future.

I am among those who think that Science has great beauty.

Marie Skłodowska-Curie (1867-1934)

1 Introduction

As currently understood, the Dirac theory [1, 2] employs a 4×1 type wavefunction, ψ. This 4×1 Dirac wavefunction is acted upon by a 4×4 Dirac Hamiltonian operator, H_{D}, in which process, four independent particle solutions result, i.e. $\psi[1]$, $\psi[2]$, $\psi[3]$, and $\psi[4]$. To this day, insofar as the real physical meaning and distinction of these four solutions, it remains unclear what these solutions really mean. We demonstrate herein that these four independent particle solutions can be brought together under a single roof wherein the Dirac wavefunction takes a new form as a 4×4 wavefunction. To that end, we shall start by introducing the well-known Dirac equation.

That is to say: for a particle whose rest-mass and wavefunction are m_0 and ψ respectively, the corresponding Dirac equation is given by:

$$\imath\hbar\gamma^{\mu}\partial_{\mu}\psi = m_0 c_0 \psi, \tag{1}$$

where: $\hbar = 1.054571817 \times 10^{-34}\,\mathrm{J\,s}$ (CODATA 2018) is the normalized Planck constant, $c_0 = 299792458 \times 10^{8}\,\mathrm{m\,s^{-1}}$ (CODATA 2018) is the speed of light in *vacuo*, $\imath = \sqrt{-1}$, and:

$$\gamma^0 = \begin{pmatrix} \mathcal{I}_2 & \emptyset \\ \emptyset & -\mathcal{I}_2 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} \emptyset & \sigma^i \\ -\sigma^i & \emptyset \end{pmatrix}, \tag{2}$$

are the 4×4 Dirac γ-matrices where $\mathcal{I}_2$ and $\emptyset$ are the 2×2 identity and null matrices respectively, and the four component Dirac wave-function, ψ, is defined as follows:

$$\psi = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}, \tag{3}$$

is the 4×1 Dirac four component wavefunction and ψ_L and ψ_R are the Dirac [1, 2] bispinors that are defined such that:

$$\psi_L = \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}, \quad \text{and,} \quad \psi_R = \begin{pmatrix} \psi_2 \\ \psi_3 \end{pmatrix}. \tag{4}$$

Throughout this paper, unless otherwise specified, the Greek indices will be understood to mean $(\mu, \nu, \ldots = 0, 1, 2, 3)$ and the lower case English alphabet indices $(i, j, k \ldots = 1, 2, 3)$.

The Dirac equation can be recast into the Schrödinger formalism as follows: $H_{\mathrm{D}}\psi = E\psi$, where $E = -\imath\hbar\partial/\partial t$ is the usual quantum mechanical energy operator, and:

$$H_{\mathrm{D}} = \imath\hbar c_0 \gamma^j \frac{\partial}{\partial x^j} - \gamma^0 m_0 c_0^2, \tag{5}$$

is the Dirac Hamiltonian operator. In §4, we shall for the purposes of efficiently making our point regarding the 4×4 wavefunction approach, use the Dirac equation in the Schrödinger formalism.

Now, in closing this introductory section, we shall give the synopsis of the present paper. In §2, we shall for instructive, completeness and self-containment purposes, present the traditional free particle solutions of the Dirac equation. Thereafter in §3, we shall discuss some of the major shortcomings of the Dirac equation – this we do in order to demonstrate that there still is a lot more about the Dirac equation that still needs to be understood. Then, in §4, we present the main task of the present paper – i.e. the Dirac wavefunction is cast into a 4×4 type wavefunction. Thereafter in §5, we proceed to make our suggestion regarding the new formulation of the Dirac wavefunction. Lastly, in §6 a general discussion is given and no conclusion is made.

2 Free particle solutions of the Dirac equation

The free particle solutions of the Dirac equation are obtained by assuming a free particle wavefunction of the form: $\psi = ue^{\imath p_\mu x^\mu/\hbar}$, where: u, is a four component object, i.e. $u^T = (u_0\ u_1\ u_2\ u_3)$, where the superscript-$T$ on u is the transpose operator. This u-function is assumed to have no space and time dependence. With this in mind, one will proceed to substituting this free particle solution: $\psi = ue^{\imath p_\mu x^\mu/\hbar}$, into (1), where-after some elementary algebraic operations – they will be led to the following linear quad-set of simultaneous equations:

$$(E - m_0c_0^2)u_0 - c_0(p_x - ip_y)u_3 - cp_zu_2 = 0, \qquad (6a)$$

$$(E - m_0c_0^2)u_1 - c_0(p_x + ip_y)u_2 + cp_zu_3 = 0, \qquad (6b)$$

$$(E + m_0c_0^2)u_2 - c_0(p_x - ip_y)u_1 - cp_zu_0 = 0, \qquad (6c)$$

$$(E + m_0c_0^2)u_3 - c_0(p_x + ip_y)u_0 + cp_zu_1 = 0. \qquad (6d)$$

An important fact to note about the above array or set of simultaneous equations is that the four solutions u_j (where: $j = 0, 1, 2, 3$) are superluminally entangled, that is to say, a change in one of the components affects every other component instantaneously i.e. in zero time interval. What this means is that for linearly dependent solutions of u_j, the Dirac equation – just as it predicts spin as a relativistic quantum phenomenon, it also predicts entanglement as a quantum phenomenon. If they exist as a separate reality in different regions of space, then, the particles, ψ_L, and, ψ_R, are entangled.

Without further ado, we shall now present the four formal solutions of the Dirac equation, these are given by: $\psi[k] = u[k]\exp\left(\imath p_\mu x^\mu/\hbar\right)$, where the $u[k]$'s are such that:

$$u[1] = \begin{pmatrix} 1 \\ 0 \\ \dfrac{c_0p_z}{E + m_0c_0^2} \\ \dfrac{c_0(p_x + ip_y)}{E + m_0c_0^2} \end{pmatrix} \quad u[2] = \begin{pmatrix} 0 \\ 1 \\ \dfrac{c_0(p_x - ip_y)}{E + m_0c_0^2} \\ -\dfrac{c_0p_z}{E + m_0c_0^2} \end{pmatrix} \qquad (7)$$

$$u[3] = \begin{pmatrix} \dfrac{c_0p_z}{E - m_0c_0^2} \\ \dfrac{c_0(p_x + ip_y)}{E - m_0c_0^2} \\ 1 \\ 0 \end{pmatrix} \quad u[4] = \begin{pmatrix} \dfrac{c_0(p_x - ip_y)}{E - m_0c_0^2} \\ -\dfrac{c_0p_z}{E - m_0c_0^2} \\ 0 \\ 1 \end{pmatrix}$$

These solutions (2) are obtained as follows:

1. From (6), u_0, and u_1 are fixed so that: $u_0 = 1$, and, $u_1 = 0$, and the resultant set of equations is solved for u_2, and, u_3.
2. Similarly, from (6), u_0, and u_1 are fixed so that: $u_0 = 0$, and, $u_1 = 1$, and the resultant set of equations is solved for u_2, and, u_3.
3. Again, from (6), u_2, and u_3 are fixed so that: $u_2 = 1$, and, $u_3 = 0$, and the resultant set of equations is solved for u_0, and, u_1.
4. Similarly, from (6), u_2, and u_3 are fixed so that: $u_2 = 0$, and, $u_3 = 1$, and the resultant set of equations is solved for u_0, and, u_1.

Now, having presented the solutions of the Dirac equation, we shall proceed to present what we feel are some of the important major shortcomings of the Dirac equation.

3 Major shortcomings of the Dirac equation

While the Dirac equation is one of the most successful equations in physics, it is not without its own shortcomings. We here briefly review some of its shortcomings.

3.1 Anomalous gyromagnetic ratios

It is a well-known fact that in its bare and natural form, the Dirac equation predicts a gyromagnetic ratio ($\mathrm{g_D}$) equal to two (i.e. $\mathrm{g_D} = 2$) and this prediction is very close to the gyromagnetic ratio of the electron [$\mathrm{g_E} = 2 + 0.002319304362(2)$], hence, the Dirac equation is said to give a good description of the electron. On the contrary, the spin-1/2 proton ($\mathrm{g_E}$) and neutron ($\mathrm{g_N}$), which – like the electron – are thought to be fundamental particles and thus are naturally expected to readily submit to a successful description by the Dirac equation – these particles have gyromagnetic ratios that are at variance i.e. ($\mathrm{g_P} = 2 + 3.5856947(5)$; $\mathrm{g_N} = 2 - 5.8260855(9)$) with the Dirac prediction. The Dirac equation lacks in its nature infrastructure the devices to correctly predict the g-ratio of any arbitrary spin-1/2 particle. This state of affairs and aspect of the Dirac equation is very disappointing. In a future paper, we will propose a solution to this problem. We must say that, in the existing literature, there exists appropriate amendments that have been made to the Dirac theory in order to solve this problem. However, these solutions lack the much needed universal character.

3.2 Negative energy solutions

Further – as is well-known, one of Dirac's purpose in the formulation of his equation was to eliminate the unwanted negative energy solutions present in the Klein-Gordon equation [3, 4]. However, negative energy solutions are still present in Dirac's equation, and this led Dirac [5] to accept these solutions as physically realistic and to propose the existence of antimatter. Carl Anderson [6] confirmed Dirac's hypothesis and latter, Giuseppe Occhialini and Patrick Blackett [7] did the same.

The existence of antimatter is now commonplace in the scientific literature. What is not clear about this antimatter particles (antiparticles) is whether or not they have negative energy and mass. Do antimatter particles fall up or down in a gravitational field? Experiments [8] are not clear and this question still needs to be answered (see *e.g.* [9–12]). For according to Einstein's [13] mass energy equivalence ($E = mc_0^2$), if the energy of antiparticles is negative, their mass should be negative too. If this is the case, it follows from *Newton's Law of Gravitation* that in a gravitational field, antiparticles aught to fall up and not down!

3.3 Whereabouts of antimatter

Furthermore – apart from the the question of whether antiparticles fall up or down, there is the yet to be an answer to the question of the whereabouts of this antimatter [14–16]. The Dirac equation not only is symmetric under electric *Charge Conjugation* (C), but, symmetric under all the known discrete symmetries of *Time* (P) and *Parity* (P) reversal including the combination of any of these these discrete symmetries – i.e. CP, CT, PT, and, CPT. This symmetric nature of the Dirac equation leads to the prediction that the Universe must contain equal amounts of matter and antimatter. This is at variance with physical experience. Otherwise, due to the annihilation of matter and antimatter into radiation, the Universe would be a radiation bath, which is clearly not the case.

This prediction of the Dirac equation is '*very unfortunate*' because it is at complete variance with physical and natural as we know and experience it. That is to say, given that matter and antimatter will annihilate to form radiation should they ever come into contact, the exist of equal positions of matter and antimatter in the Universe would mean that if the Dirac prediction on the matter-antimatter census is correct, then, the Universe aught to be no more than a radiation bath. Clearly, this is not what we see around us.

3.4 Lack of a universal character

Additionally – every fermion particle (electron, proton, neutron, neutrinos, quarks *etc*) in particle physics is described by the Dirac equation. This gives the superficial impression that the Dirac equation is a universal equation for all spin-1/2 particles. A closer look will reveal that, while this equation is used to describe fermion, it needs to be supplemented in order to match-up with experimental data. As already pointed out in §3.1, the g-ratios of every other particle save for that of the electron are not in conformity with the natural Dirac equation. If the Dirac equation were indeed a universal equation for all fermions, it must contain within its natural infrastructure the necessary adjustable parameters that would make it fit with the experimental data of a given particle. These post-experimental adjustments that are made in order that the Dirac equation fits to experimental data are of *ad hoc* nature.

Apart from the inability to explain in a smooth manner the g-ratios of different fermions, we have the issue of the universality of spin. That is to say, the Dirac equation is an equation only capable of explaining spin-1/2 particles, an not any general spin particle. For example, in order to explain spin-3/2, we need to find another equation for this – the Rarita-Schwinger equation [17] in this case. In general, fermions have spin $\pm n/2$ with: $n = 1, 3, 5, 7, ..., 2r + 1, etc$. This implies that a new equation is required for every spin particle.

3.5 Fundamental origins of the Dirac equation

Lastly – another very important and yet largely ignored reality is that of the fundamental origins of the Dirac equation. That is to say, despite success, it remains that Dirac guessed his equation – albeit in a very educated manner. All he sought was an equation linear in both the space and time derivatives such that when this equation is "squared" it would yield the Klein-Gordon equation. It can be said that this issue of the origins of the Dirac equation is not unique to the Dirac equation, but all quantum mechanical equations.

The Klein-Gordon equation is derived from the well-known Einstein [18] energy-momentum dispersion relation: $E^2 = p^2c_0^2 + m_0^2c_0^4$, *via* the successful method of canonical quantization that was used by Schrödinger to arrive at his successful equation that describes the atomic world. Dirac's method of arriving at his equation is not fundamental at all, and to this day, no real progress on this has been made. Where does the Dirac equation really come from? This is yet another question that also needs an answer.

4 4 × 4 Dirac wavefunction

The very fact that the Dirac Hamiltonian, H_D, is a 4×4 component object acting on, ψ, this readily implies that, ψ, can be a $4 \times k$ component object where: $k = 1, 2, 3, 4, 5, etc$. If: $1 \leq k < 4$, the resulting system of equations is overdetermined and will thus have more than one solution, and if: $k > 4$, the resultant system of equations is under-determined and is unable to yield a solution. If: $k = 4$, the system has one and only solution, and this is the case of the 4×4 Dirac wavefunction that we would like to have a look at.

In the event of a 4×4 Dirac wavefunction where as usual:

$\psi = ue^{\imath p_\mu x^\mu/\hbar}$, the u-function is such that:

$$u = \begin{pmatrix} u_{00} & u_{01} & u_{02} & u_{03} \\ u_{10} & u_{11} & u_{12} & u_{13} \\ u_{20} & u_{21} & u_{22} & u_{23} \\ u_{30} & u_{31} & u_{32} & u_{33} \end{pmatrix} = \begin{pmatrix} u_a & u_b \\ u_c & u_d \end{pmatrix}, \tag{8}$$

where likewise:

$$\begin{matrix} u_a = \begin{pmatrix} u_{00} & u_{01} \\ u_{10} & u_{11} \end{pmatrix} & u_b = \begin{pmatrix} u_{02} & u_{03} \\ u_{12} & u_{13} \end{pmatrix} \\ u_c = \begin{pmatrix} u_{20} & u_{21} \\ u_{30} & u_{31} \end{pmatrix} & u_d = \begin{pmatrix} u_{22} & u_{23} \\ u_{32} & u_{33} \end{pmatrix} \end{matrix}. \tag{9}$$

Of this 4×4 component wavefunction, ψ, we shall require of it to observe the following constraint:

$$\psi^\dagger \psi = \psi \psi^\dagger = \varrho \mathcal{I}_4, \tag{10}$$

where: $\mathcal{I}_4$, is the 4×4 identity matrix, and, $\varrho \in \mathbb{R}$, is a real zero-rank object – it is the quantum mechanical probability density amplitude. This constraint i.e. (10) is required by the unified theory of gravitation and electromagnetism [19] that we are currently working on.

Now, substituting the new 4×4 component wavefunction into the Dirac equation (1), we will have:

$$\begin{pmatrix} (E - m_0 c_0^2)\mathcal{I}_2 & -c_0 \vec{\sigma} \cdot \vec{p} \\ c_0 \vec{\sigma} \cdot \vec{p} & -(E + m_0 c_0^2)\mathcal{I}_2 \end{pmatrix} \begin{pmatrix} u_a & u_b \\ u_c & u_d \end{pmatrix} = 0. \tag{11}$$

As we proceed, the reader must take note of the fact that the object, $\vec{\sigma} \cdot \vec{p}$, is a 2×2 matrix, i.e.:

$$\vec{\sigma} \cdot \vec{p} = \begin{pmatrix} p_z & p_x - \imath p_y \\ p_x + \imath p_y & -p_z \end{pmatrix}. \tag{12}$$

This matrix, $\vec{\sigma} \cdot \vec{p}$, is hermitian.

Now, from (11), four equations will result and these are:

$$u_a = \left(\frac{c_0 \vec{\sigma} \cdot \vec{p}}{E - m_0 c_0^2} \right) u_c, \tag{13a}$$

$$u_b = \left(\frac{c_0 \vec{\sigma} \cdot \vec{p}}{E - m_0 c_0^2} \right) u_d, \tag{13b}$$

$$u_c = \left(\frac{c_0 \vec{\sigma} \cdot \vec{p}}{E + m_0 c_0^2} \right) u_a, \tag{13c}$$

$$u_d = \left(\frac{c_0 \vec{\sigma} \cdot \vec{p}}{E + m_0 c_0^2} \right) u_b. \tag{13d}$$

For a solution to this set of simultaneous equation, we shall set as a constraint the following:

$$u_a = u_d = \mathcal{I}_2 \sqrt{\varrho/2}. \tag{14}$$

This naturally leads to the following for, u_b, and, u_c, i.e.:

$$u_b = \sqrt{\varrho/2} \left(\frac{c_0 \vec{\sigma} \cdot \vec{p}}{E - m_0 c_0^2} \right). \tag{15a}$$

$$u_c = \sqrt{\varrho/2} \left(\frac{c_0 \vec{\sigma} \cdot \vec{p}}{E + m_0 c_0^2} \right). \tag{15b}$$

Hence:

$$u = \sqrt{\varrho/2} \begin{pmatrix} \mathcal{I}_2 & \dfrac{c_0 \vec{\sigma} \cdot \vec{p}}{E - m_0 c_0^2} \\ \dfrac{c_0 \vec{\sigma} \cdot \vec{p}}{E + m_0 c_0^2} & \mathcal{I}_2 \end{pmatrix}. \tag{16}$$

Writing this 4×4 matrix (16) in full, it will be as it appears in (17). Immediately, one will be quick to notice that columns (1), (2), (3), and (4) of this matrix (17) are in-fact the traditional solutions ($u[1], u[2], u[3], u[4]$) to the Dirac equation given in (2). What this means is that the 4×4 wavefunction is a grand synthesis of these four traditional solutions into one giant set of mutually dependent quadruplet system of particles.

5 Quarks

Apart from the simplification of bringing four independent particle solutions into a single particle solution, we suggest that this recasting of the Dirac wavefunction into a 4×4 wavefunction provides additional physical simplification in the analysis of the solution. To that end, let us start-off by writing down the full 4×4 Dirac wavefunction: $\psi = ue^{\imath p_\mu x^\mu/\hbar}$. For the 4×4 Dirac wavefunction, the u-function has been defined in (17) and from that definition, it follows that:

$$\psi = \begin{pmatrix} \psi_N & \psi_R \\ \psi_L & \psi_N \end{pmatrix}, \tag{18}$$

where – accordingly:

$$\psi_N = \mathcal{I}_2 \sqrt{\varrho/2} \exp\left(\frac{\imath p_\mu x^\mu}{\hbar} \right), \tag{19a}$$

$$\psi_R = \sqrt{\varrho/2} \left(\frac{c_0 \vec{\sigma} \cdot \vec{p}}{E - m_0 c_0^2} \right) \exp\left(\frac{\imath p_\mu x^\mu}{\hbar} \right), \tag{19b}$$

$$\psi_L = \sqrt{\varrho/2} \left(\frac{c_0 \vec{\sigma} \cdot \vec{p}}{E + m_0 c_0^2} \right) \exp\left(\frac{\imath p_\mu x^\mu}{\hbar} \right). \tag{19c}$$

In-comparison, i.e. between ψ as defined in (3) and the resultant definition of it in (18), we see that the initially four particles: ψ_a, ψ_b, ψ_c, and, ψ_d, have been reduced to three because, ψ_a, and, ψ_d, are identical – i.e. $\psi_a = \psi_d = \psi_N$. In (18), we have according to the parlance of the Dirac formalism identified ψ_b, and, ψ_c, with the right and the left-handed Dirac components. In terms of handedness, we have in the same parlance defined a new form of handedness in the ψ_N-particle,

$$u = \sqrt{\varrho/2}\begin{pmatrix} 1 & 0 & \dfrac{c_0 p_z}{E - m_0 c_0^2} & \dfrac{c_0(p_x - \imath p_y)}{E - m_0 c_0^2} \\ 0 & 1 & \dfrac{c_0(p_x + \imath p_y)}{E - m_0 c_0^2} & -\dfrac{c_0 p_z}{E - m_0 c_0^2} \\ \dfrac{c_0 p_z}{E + m_0 c_0^2} & \dfrac{c_0(p_x - \imath p_y)}{E + m_0 c_0^2} & 1 & 0 \\ \dfrac{c_0(p_x + \imath p_y)}{E + m_0 c_0^2} & -\dfrac{c_0 p_z}{E + m_0 c_0^2} & 0 & 1 \end{pmatrix} \tag{17}$$

a handedness that we shall call – *neutral-handedness*, hence, ψ_N, is a neutral-handed particle, this particle is neither left nor right-handed, hence our calling it neutral-handed particle and hence the subscript-N in its denotation.

Now, in the set: ψ_N, ψ_R, and, ψ_L, we have a trio of particles that are not only mutually dependent but entangled, and in addition to this, they are confined in the same region of space. Each of these particles do not exist independently of the other, they can never be free of each other far-away from the region defined by the ψ-particle system. The boundary in spacetime of the ψ-particle system is defined by the normalization of conditions of this particle system, i.e. $\langle \psi \,||\, \psi \rangle = \mathcal{I}_4$.

Now, given the following there facts:

1. The proton and neutron are each known to contain three quarks living inside them.
2. Further, the quarks strongly appear to be unable to exist independent of each other.
3. Furthermore, these quarks strongly appear to be eternal prisoners inside the proton and neutron. They are unable to exist beyond the radius demarcating the proton and neutron particle systems.

From these facts – i.e. the obvious similarity in the nature of quarks and the trio ψ_N, ψ_R, ψ_L, it is natural to wonder whether or not these three particles ψ_N, ψ_R, ψ_L are the quarks whose origins we have thus far elusively sought to understand? From this viewpoint, the present recasting of the Dirac wavefunction surely opens up a new avenue of thinking regarding the Dirac equation and quarks, hence justifying the need to seriously consider the 4×4 Dirac wavefunction. With that having been said, we must at this very juncture say that – it is not our intention to explore this idea that the set (ψ_N, ψ_R, ψ_L) might explain quarks and the reason for this is simple that we feel it is too early for us to do so, otherwise all that we would do is to speculate.

6 Discussion

As currently accepted and understood, the Dirac theory [1, 2] employs a 4 × 1 type wavefunction. This 4 × 1 Dirac wavefunction is acted upon by a 4×4 Dirac Hamiltonian, in which process, four independent particle solutions result and insofar as the real physical meaning and distinction of these four solutions, it is not clear what these solutions really mean. It is this that this paper has made an endeavour to provoke a thought process were a physical meaning can be attached to these four independent particle solutions of the Dirac equation and this is *via* the recasting of the Dirac wavefunction into a 4 × 4 type wavefunction.

We first presented this idea of a 4 × 4 Dirac wavefunction in [20, 21]. Prior to the said presentation [20, 21], we had never seen or heard of it anywhere in the literature. Therein [20], this idea was presented as no more than a *mathematical curiosity*, with no physical meaning attached to it. We had to come back to this idea now because we realised that it is necessary for the theory that we are currently working on [19], that is, a unified field theory of the gravitational and electromagnetic phenomenon.

What we have herein done with Dirac's four independent particle solutions, is to demonstrate that these can be represented as a quadruplet particle system wherein the Dirac wavefunction takes a new form as a 4×4 wavefunction. In this new formation, these four particle solutions precipitate into three distinct and mutuality dependent particles (ψ_L, ψ_N, ψ_R) that are permanently bound in the same region of space.

Realizing that the proton and neutron are composite particles each comprising three quarks that are in (color) confinement, we proceeded logically to make the natural suggestion to the effect that these Dirac particles (ψ_L, ψ_N, ψ_R) might be quarks. Whether or not these particles are quarks, this surely is something that further investigations will have to be establish.

Received on March 6, 2021

References

1. Dirac P. A. M. The Quantum Theory of the Electron. *Proc. Roy. Soc. Lon. A: Math. Phys. & Eng. Sci.*, 1928, v. 117 (778), 610–624.
2. Dirac P. A. M. The Quantum Theory of the Electron. Part II. *Proc. Roy. Soc. Lon. A: Math., Phys. & Eng. Sci.*, 1928, v. 118 (779), 351–361.
3. Klein O. Quantentheorie und fünfdimensionale Relativitätstheorie. *Zeitschrift für Physik*, 1926, v. 37 (12), 895–906.
4. Gordon W. Der Comptoneffekt nach der Schrödingerschen Theorie. *Zeitschrift für Physik*, 1926, v. 40 (1), 117–133.
5. Dirac P. A. M. A Theory of Electrons and Protons. *Proc. Roy. Soc. Lon. A: Math., Phys. & Eng. Sci.*, 1930, v. 126 (801), 360–365.

6. Anderson C. A. The Positive Electron. *Phys. Rev.*, 1933, v. 43 (6), 491–494.
7. Fitch V. L. Elementary Particle Physics: The Origins. *Rev. Mod. Phys.*, 1999, v. 71 (2), S25–S32.
8. Charman A. E., Amole C., Ashkezari M. D., *at al*. Description and First Application of a New technique to Measure the Gravitational Mass of Antihydrogen. *Nature Comm.*, 2013, v. 4 (1).
9. Schiff L. I. Sign of the Gravitational Mass of a Positron. *Phys. Rev. Lett.*, 1958, v. 1 (7), 254–255.
10. Morrison P. Approximate Nature of Physical Symmetries. *Am. J. Phys.*, 1958, v. 26 (6), 358–368.
11. Nieto M. M. and Goldman T. The Arguments Against Antigravity and the Gravitational Acceleration of Antimatter. *Phys. Rep.*, 1991, v. 205 (5), 221–281.
12. Apostolakis A., Aslanides E., Backenstoss G. , *et al.* Tests of the Equivalence Principle with Neutral Kaons. *Phys. Lett. B*, 1999, v. 452 (3-4), 425–433.
13. Einstein A. Ist die Trägheit eines Körpers von Seinem Energieinhalt Abhängig? *Ann. der Phys.*, 1905, v. 323 (13), 639–641.
14. Sakhorov A. D. Violation of CP Symmetry, C-Asymmetry and Baryon Asymmetry of the Universe. *Pisma Zh. Eksp. Teor. Fiz.*, 1967, v. 5, 32–35.
15. Weidenspointner G., Skinner G., Jean P., and *et al.* An Asymmetric Distribution of Positrons in the Galactic disk Revealed by γ-Rays. *Nature*, 2008, v. 451 (7175), 159–162.
16. Neri N. and The LHCb collaboration. Measurement of Matter-Antimatter Differences in Beauty Baryon Decays. *Nature Phys.*, 2017, v. 13 (4), 391–396.
17. Rarita W. and Schwinger J. S. On a Theory of Particles with Half-Integral Spin. *Phys. Rev.*,1941, v. 60 (1), 61–61.
18. Einstein A. Zur Elektrodynamik bewegter Körper (On the Electrodynamics of Moving Bodies). *Ann. der Phys.*, 1905, v. 322 (10), 891–921.
19. Nyambuya G. G. Fundamental Geometrodynamic Justification of Gravitomagnetism (I). *Prog. Phys.*, 2020, v. 16 (2), 73–91.
20. Nyambuya G. G. On the Dirac Wavefunction as a 4×4 Component Function. *Prespacetime Journal*, 2016, v. 7 (8), 1232–1243.
21. Nyambuya G. G. On the Possible Origins of the Dirac Equation. *Prespacetime Journal*, 2016, v. 7 (13), 1818–1826.

The Exact Gödel Metric

Patrick Marquet
Calais, France. E-mail: patrick.marquet6@wanadoo.fr

We demonstrate that Gödel's metric does not represent a model of universe as it is usually accepted in the standard literature. In fact, a close inspection shows that this metric as it stands is a very special case of a broader metric. Introducing a simple conformal transformation readily induces a pressure term on the right hand side of the Einstein's field equations which actually describe a peculiar perfect fluid. This term was wrongly interpreted by Gödel as the *ad hoc* cosmological constant required to sustain his model. Gödel's space-time can be thus regarded as a real physical system with no cosmological implication and it is relegated to the class of ordinary metrics. The emergence of the related closed time-like curves is not bound to a rotating universe as stated in all classical treatments and this fact naturally sheds new light on time travel feasibility considerations.

Notations

Space-time greek indices run from: α,β: 0, 1, 2, 3.
Space-time signature: -2.
$\varkappa$ is the Einstein constant.
We adopt here: $c = 1$.

1 The Gödel universe

1.1 General

In his original paper [1], Kurt Gödel has derived an exact solution to Einstein's field equations in which the matter takes the form of a shear/pressure free fluid (dust solution).

This universe is homogeneous but non-isotropic and it exhibits a specific rotational symmetry which allows for the existence of *close timelike curves* (*CTCs*). The Gödel space-time has a five dimensional group of isometries (G5) which is transitive. (An action of a group is transitive on a manifold (M,$\boldsymbol{g}$), if it can map any point of the manifold into any other points of M).

It admits a five dimensional *Lie algebra* of *Killing vector fields* generated by a time translation ∂_{x_0}, two spatial translations ∂_{x_1}, ∂_{x_2}, plus two further Killing vector fields:

$$\partial_{x_3} - x_2\partial_{x_2} \quad \text{and} \quad 2\mathrm{e}^{x_1}\partial_{x_0} + x_2\partial_{x_3} + \left(\mathrm{e}^{2x_1} - \frac{1}{2}\, {x_2}^2\partial_{x_2}\right).$$

The *Weyl tensor* of the *standard* Gödel solution has Petrov type D:

$$C^{\alpha\beta}{}_{\mu\nu} = R^{\alpha\beta}{}_{\mu\nu} + \frac{R}{3}\,\delta^{\alpha}{}_{[\mu}\delta_{\nu]}{}^{\beta} + 2\delta^{[\alpha}{}_{[\mu}R_{\nu]}{}^{\beta]}.$$

The presence of the non-vanishing Weyl tensor prevents the Gödel metric from being *Euclidean* whatever the coordinates transformations.

This is in contrast to the Friedmann-Lemaître-Robertson-Walker metric which can be shown to reduce to a conformal Euclidean metric, implying that its Weyl tensor is zero [2].

The Gödel universe is often dismissed because it implies a non zero cosmological term and also since its rotation would conflict with observational data.

In what follows, we are able to relax our demand that the Gödel metric be a description of an actual universe. This is achieved through a specific transformation which makes Gödel space-time an "ordinary" metric just as any other metrics currently derived in physics.

1.2 The basic theory

The classical Gödel line element is generically given by the interval

$$ds^2 = a^2\left[{dx_0}^2 - {dx_1}^2 + {dx_2}^2\frac{1}{2}\,\mathrm{e}^{2x_1} - {dx_3}^2 + 2\mathrm{e}^{x_1}(dx_0dx_2)\right], \qquad (1.1)$$

or equivalently:

$$ds^2 = a^2\left[-{dx_1}^2 - {dx_3}^2 - {dx_2}^2\frac{1}{2}\,\mathrm{e}^{2x_1} + (\mathrm{e}^{x_1}\,dx_2 + dx_0)^2\right]. \qquad (1.2)$$

$a > 0$ is a constant.

The components of the metric tensor are:

$$(g_{\mu\nu})_{\mathrm{G}} = \begin{pmatrix} a^2 & 0 & a^2\mathrm{e}^{x_1} & 0 \\ 0 & -a^2 & 0 & 0 \\ a^2\mathrm{e}^{x_1} & 0 & a^2\frac{1}{2}\mathrm{e}^{2x_1} & 0 \\ 0 & 0 & 0 & -a^2 \end{pmatrix},$$

$$(g^{\mu\nu})_{\mathrm{G}} = \begin{pmatrix} -a^2 & 0 & -a^{-2}2\mathrm{e}^{-x_1} & 0 \\ 0 & -a^2 & 0 & 0 \\ -a^{-2}2\mathrm{e}^{-x_1} & 0 & -a^{-2}2\mathrm{e}^{-2x_1} & 0 \\ 0 & 0 & 0 & -a^2 \end{pmatrix}.$$

In this particular case, since only $\partial_1(g_{22})_{\mathrm{G}} \neq 0$ and $\partial_1(g_{02})_{\mathrm{G}} \neq 0$, the non-zero connection components are:

$$\Gamma^0{}_{01} = 1 \qquad \Gamma^0{}_{12} = \Gamma^1{}_{02} = \frac{1}{2}\,\mathrm{e}^{x_1}$$

$$\Gamma^1{}_{22} = \frac{1}{2}\,\mathrm{e}^{2x_1} \qquad \Gamma^2{}_{01} = -\mathrm{e}^{-x_1}$$

Those greatly simplify the Ricci tensor: $R_{\beta\gamma} = \partial_1\Gamma^1{}_{\beta\gamma} + \Gamma^1{}_{\beta\gamma} - \Gamma^{\delta}{}_{\alpha\beta}\Gamma^{\alpha}{}_{\delta\gamma}$ whose components reduce to:

$$R_{00} = 1, \qquad R_{22} = \mathrm{e}^{2x_1}$$

$$R_{02} = R_{20} = \mathrm{e}^{x_1}$$

The Gödel *unit vector* $\boldsymbol{u}$ of matter in the direction of the x_0 lines has the following components:

$$(u^\mu)_{\mathrm{G}} = (a^{-1}, 0, 0, 0), \tag{1.3}$$

$$\left(u_\mu\right)_{\mathrm{G}} = (a, 0, a\mathrm{e}^{x_1}, 0), \tag{1.4}$$

hence:

$$R_{\mu\nu} = \left(u_\mu u_\nu\right)_{\mathrm{G}} a^{-2}, \tag{1.5}$$

$$R = \left(u^\mu u_\mu\right)_{\mathrm{G}} = a^{-2}. \tag{1.6}$$

In order to make his metric a compatible solution to Einstein's field equations, Gödel is led to introduce a cosmological constant Λ as:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \varkappa\rho u_\mu u_\nu + g_{\mu\nu}\Lambda. \tag{1.7}$$

To achieve this compatibility, he then further sets:

$$a^{-2} = \varkappa\rho, \tag{1.8}$$

$$\Lambda = -\frac{1}{2}R = -\frac{1}{2a^2} = -\frac{1}{2}\varkappa\rho. \tag{1.9}$$

As primarily claimed by Gödel, its stationary space-time is homogeneous.

For every point A of the manifold $(\mathrm{M}, \boldsymbol{g}_{\mathrm{G}})$, there exists a one-parameter group of transformations of M carrying A into itself.

This means that $(\mathrm{M}, \boldsymbol{g}_{\mathrm{G}})$ has a rotational symmetry and matter rotates everywhere with a constant rotation velocity magnitude $\boldsymbol{\omega}_{\mathrm{G}}$ orthogonal to $\boldsymbol{u}_{\mathrm{G}}$.

Using the contravariant components:

$$(\omega^\alpha)_{\mathrm{G}} = \left(0, 0, 0, \frac{\sqrt{2}}{a^2}\right), \tag{1.10}$$

one finds:

$$\omega_{\mathrm{G}} = (g_{\alpha\mu}\omega^\alpha\omega^\mu)_{\mathrm{G}}{}^{1/2} = \frac{a}{\sqrt{2}}. \tag{1.11}$$

With (1.8) this magnitude is:

$$\boldsymbol{\omega}_{\mathrm{G}} = \left(\frac{1}{2}\varkappa\rho\right)^{1/2}. \tag{1.12}$$

A first glance at these constraints, readily reveals a fairly high degree of arbitrariness in the theory.

Finetuning the hypothetical constant Λ with the density of the universe (and the Ricci scalar) appears indeed as a somewhat dubious physical argument.

We shall see that those ill-defined assumptions are not required in order for the basic model to satisfy the field equations.

2 Gödel's model defined as a homogenous perfect fluid

2.1 Reformulation of Gödel's metric

We now make the assumption that a is slightly space-time variable and we set:

$$a^2 = \mathrm{e}^{2U}. \tag{2.1}$$

The positive scalar $U(x_\mu) > 0$ will be explicited below.

The Gödel metric tensor components (1.2) are related to the fundamental metric tensor $\boldsymbol{g}$ by:

$$(g_{\mu\nu})_{\mathrm{G}} = \mathrm{e}^{2U} g_{\mu\nu}, \tag{2.2}$$

$$(g^{\mu\nu})_{\mathrm{G}} = \mathrm{e}^{-2U} g^{\mu\nu}, \tag{2.2 bis}$$

This means that the Gödel metric is now conformal:

$$ds^2 = \mathrm{e}^{2U}\left[dx_0{}^2 - dx_1{}^2 + dx_2{}^2\frac{1}{2}\,\mathrm{e}^{2x_1} - dx_3{}^2 + 2\mathrm{e}^{x_1}(dx_0 dx_2)\right]. \tag{2.3}$$

We are now going to see how the substitution (2.1) drastically changes the meaning of Gödel's limited theory.

2.2 Relativistic analysis of a neutral homogeneous perfect fluid

2.2.1 The geodesic differential system

Let us consider the manifold $(\mathrm{M}, \boldsymbol{g})$, on which is defined a vector tangent to the curve C in local coordinates:

$$\dot{x}^\alpha = \frac{dx^\alpha}{d\zeta}, \quad \text{where } \zeta \text{ is an affine parameter.}$$

In these coordinates we consider the scalar valued function $f(x^\alpha, \dot{x}^\alpha)$ which is homogeneous and of first degree with respect to $\dot{x}^\alpha$.

To the curve C joining the point x_1 to x_2, one can always associate the integral $\mathcal{A}$ such that

$$\mathcal{A} = \int_{\zeta_1}^{\zeta_2} f(x^\alpha, \dot{x}^\alpha) d\zeta = \int_{x_1}^{x_2} f(x^\alpha, \dot{x}^\alpha) dx^\alpha. \tag{2.4}$$

We now want to evaluate the variation of $\mathcal{A}$ with respect to the points ζ_1 and ζ_2:

$$\delta\mathcal{A} = f\delta_{\zeta_2} - f\delta_{\zeta_1} - \int_{\zeta_1}^{\zeta_2} \delta f d\zeta.$$

Classically we know that:

$$\int_{\zeta_1}^{\zeta_2} \delta f d\zeta = \left[\frac{\partial f}{\partial \dot{x}^\alpha}\delta x^\alpha\right] - \int_{\zeta_1}^{\zeta_2} E_\alpha \delta x^\alpha d\zeta,$$

where E_α is the first member of the *Euler equations* associated with the function f.

With E_α as the components of $\boldsymbol{E}$, we infer the expression

$$\delta\mathcal{A} = [w(\delta)]_{x_2} - [w(\delta)]_{x_1} - \int_{\zeta_1}^{\zeta_2} \boldsymbol{E}\delta x d\zeta, \tag{2.5}$$

where $[w(\delta)]$ has the form:

$$[w(\delta)] = \left(\frac{\partial f}{\partial \dot{x}^\alpha}\right)\delta x^\alpha - \left(x^\alpha \frac{\partial f}{\partial \dot{x}^\alpha} - f\right)\delta\zeta .$$

Due to the homogeneity of f, it reduces to:

$$w(\delta) = \left(\frac{\partial f}{\partial \dot{x}^\alpha}\right)\delta x^\alpha .$$

Let us apply the above results to the function

$$f = \mathrm{e}^U \frac{ds}{d\zeta} = \mathrm{e}^U \left(g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta\right)^{1/2},$$

where e^U is defined everywhere on $(\mathrm{M},\boldsymbol{g})$.

Between two points x_1 and x_2, of $(\mathrm{M},\boldsymbol{g})$ connected by a time-like curve we have the correspondence:

$$s' = \int_{x_1}^{x_2} \mathrm{e}^U ds = \int_{x_1}^{x_2} \mathrm{e}^U \left(g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta\right)^{1/2}. \tag{2.6}$$

We first differentiate $f^2 = \mathrm{e}^{2U}\left(g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta\right)$ with respect to $\dot{x}^\alpha$ and x^α:

$$f \frac{\partial f}{\partial \dot{x}^\alpha} = \mathrm{e}^{2U} g_{\alpha\beta}\dot{x}^\beta , \tag{2.7}$$

$$f \frac{\partial f}{\partial x^\alpha} = \mathrm{e}^U \left(g_{\beta\mu}\dot{x}^\beta \dot{x}^\mu\right)^{1/2} \times \left[\partial_\alpha \mathrm{e}^U \left(g_{\beta\mu}\dot{x}^\beta \dot{x}^\mu\right)^{1/2} + \frac{1}{2}\,\mathrm{e}^U \partial_\alpha \left(g_{\beta\mu}\dot{x}^\beta \dot{x}^\mu\right)\right]. \tag{2.8}$$

We now choose s as the affine parameter ζ of the curve C, so the vector $\dot{x}^\beta$ is here regarded as the unit vector u^β tangent to C.

Equations (2.7) and (2.8) then reduce to the following:

$$\frac{df}{d\dot{x}^\beta} = \mathrm{e}^U u_\beta , \tag{2.9}$$

$$\frac{df}{dx^\beta} = \partial_\beta \mathrm{e}^U + \frac{1}{2}\,\mathrm{e}^U \partial_\beta (g_{\alpha\mu}) u^\alpha u^\mu ,$$

$$\frac{df}{dx^\beta} = \partial_\beta \mathrm{e}^U + \mathrm{e}^U \Gamma_{\alpha\beta,\mu} u^\alpha u^\mu . \tag{2.10}$$

The $\Gamma_{\alpha\beta,\mu}$ are here the Christoffel symbols of the first kind.

Expliciting the Euler equations $f(x^\alpha, du^\alpha)$:

$$E_\beta = \frac{d}{ds}\frac{\partial f}{\partial u^\beta} - \frac{\partial f}{\partial x^\beta}, \tag{2.11}$$

we get:

$$E_\beta = \frac{d}{ds}\left(\mathrm{e}^U u_\beta\right) - \mathrm{e}^U\left(\Gamma_{\alpha\beta,\mu} u^\alpha u^\mu\right) - \partial_\beta \mathrm{e}^U ,$$

$$E_\beta = \mathrm{e}^U\left(u^\mu \partial_\mu u_\beta - \Gamma_{\alpha\beta,\mu} u^\alpha u^\mu\right) - \partial_\alpha \mathrm{e}^U \left(\delta_\beta{}^\alpha - u^\alpha u_\beta\right),$$

$$E_\beta = \mathrm{e}^U\left[\left(u^\mu \nabla_\mu u_\beta\right) - \partial_\beta U - \partial_\alpha U \left(\delta_\beta{}^\alpha - u^\alpha u_\beta\right)\right]. \tag{2.12}$$

Equation (2.5) becomes:

$$\delta\mathcal{A} = [w(\delta)]_{x_2} - [w(\delta)]_{x_1} - \int_{x_1}^{x_2} \langle \boldsymbol{E}\delta x\rangle ds, \tag{2.13}$$

where locally: $w(\delta) = \mathrm{e}^U u_\alpha dx^\alpha$.

When the curve C varies between two *fixed* points x_1 and x_2 the local variations $[w(\delta)]_{x_2}$ and $[w(\delta)]_{x_1}$ vanish. Therefore applying the variational principle to (2.13) simply leads to:

$$\delta\mathcal{A} = -\int_{x_1}^{x_2} \langle \boldsymbol{E}\delta x\rangle ds = 0, \tag{2.14}$$

from which we infer $\boldsymbol{E} = 0$, i.e., from (2.12):

$$u^\mu \nabla_\mu u_\beta - \partial_\alpha U\left(\delta_\beta{}^\alpha - u^\alpha u_\beta\right) = 0 \qquad (\text{since } \mathrm{e}^U \neq 0). \tag{2.15}$$

The equation (2.15) is formally identical to the differential system obeyed by the flow lines of a perfect fluid of density ρ with an equation of state $\rho = f(P)$ (*see Appendix*):

$$T_{\mu\beta} = (\rho + P) u_\mu u_\beta - P g_{\mu\beta}. \tag{2.16}$$

These flow lines are thus timelike geodesics of the conformal metric to $(\mathrm{M},\boldsymbol{g})$ according to (2.6):

$$s' = \int_{s_1}^{s_2} \mathrm{e}^U ds, \tag{2.17}$$

with

$$U = \int_{P_1}^{P_2} \frac{dP}{\rho + P}. \tag{2.18}$$

All along the curve segment (s'), the pressure is varying between two endpoints s_1 and s_2 which correspond to the values P_1 and P_2.

One can find similar conclusions in [3, 4].

The positive scalar e^U accounts for the *relativistic fluid index* [5].

2.2.2 The Gödel interpretation

The tensor (2.16) can be equivalently written:

$$T_{\mu\beta} = \rho u_\mu u_\beta - P h_{\mu\beta}, \tag{2.19}$$

with the *projection tensor*:

$$h_{\mu\beta} = g_{\mu\beta} - u_\mu u_\beta . \tag{2.20}$$

The cosmological term can then be re-introduced by setting

$$P = -\frac{\Lambda}{\varkappa}, \tag{2.21}$$

yielding the model which Gödel simply focused on.

Finally, by letting a be a *conformal factor*, we see that Gödel's metric (2.3) is simply the solution of the field equations with a variable pressure term as per:

$$R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu} R = \varkappa\left(\rho u_\mu u_\beta - P h_{\mu\beta}\right). \tag{2.21bis}$$

The cosmological "constant" Λ is thus no longer this arbitrary ingredient required to sustain the Gödel model and so are the constraints (1.8) and (1.9).

2.3 The Gödel rotation

2.3.1 Vorticity of the fluid

We just showed that Gödel space-time should be likened to a perfect fluid.

The time-like 4-vector u_α is everywhere tangent to the flow lines of this fluid.

The covariant derivative $u_{\alpha;\mu}$ may be expressed in a invariant manner in terms of tensor fields which describe the kinematics of the congruence of curves generated by u^α.

In Gödel's case, the *shear tensor* $\sigma_{\alpha\mu}$ vanishes:

$$\sigma_{\alpha\mu} = u_{(\alpha;\mu)} - \frac{1}{3}\,\theta h_{\alpha\mu} + \mathring{u}_{(\alpha}u_{\mu)} = 0, \tag{2.22}$$

where θ is the *expansion scalar*:

$$\theta = u^\alpha{}_{;\alpha}. \tag{2.22bis}$$

$\mathring{u}_\alpha$ is the *acceleration vector* of the flow lines:

$$\mathring{u}_\alpha = u_{\alpha;\mu}u^\mu. \tag{2.22ter}$$

For a perfect fluid, this acceleration is shown to be (*see Appendix*):

$$\mathring{u}_\alpha = \partial_\alpha U. \tag{2.23}$$

Besides $\mathring{u}_\alpha$ and θ, the shearless fluid is characterized by the *vorticity tensor*:

$$\omega_{\alpha\mu} = h_\alpha{}^\sigma h_\mu{}^\nu u_{[\sigma;\nu]} = u_{[\alpha;\mu]} + \mathring{u}_{[\alpha}u_{\mu]}, \tag{2.24}$$

from which is derived the vorticity *4-vector* $\boldsymbol{\omega}$ of the flow lines of the fluid.

The $\boldsymbol{\omega}$-components are known to be: [6]

$$\omega^\beta = \frac{1}{2}\,\eta^{\beta\gamma\sigma\rho}u_\gamma\omega_{\sigma\rho}, \tag{2.25}$$

with the Levi-Civita tensor: $\eta^{\beta\gamma\sigma\rho} = -g^{-1/2}\cdot\varepsilon^{\beta\gamma\sigma\rho}$.

The kinematic quantities $\omega_{\alpha\mu}$ and ω_α are completely orthogonal to u^μ , i.e.,

$$\omega_{\alpha\mu}u^\mu = 0, \quad \mathring{u}_\alpha u^\mu = \omega_\alpha u^\mu = 0.$$

(Shear free flows of a perfect fluid associated with the Weyl tensor have been extensively investigated by A. Barnes, Classical General Relativity. proc. Cambridge, 1984).

2.3.2 Conformal transformations

All above results can be easily extended to the conformal manifold (M,$\boldsymbol{g}'$) by applying the covariant derivative $(\nabla_\mu)'$ formed with the *conformal connection coefficients*:

$$\left(\Gamma^\gamma{}_{\alpha\beta}\right)' = \Gamma^\gamma{}_{\alpha\beta} + 2\delta^\gamma{}_{(\alpha}U_{,\beta)} - g_{\alpha\beta}U^{,\gamma}. \tag{2.26}$$

One also defines the unit 4-vector $\boldsymbol{w}$ of the fluid on the conformal metric $(ds^2)'$ as:

$$w^\mu = \mathrm{e}^U u^\mu, \tag{2.27}$$

$$w_\beta = \mathrm{e}^{-U}u_\beta. \tag{2.28}$$

In this case, the differential system of the flow lines w^μ admits the *relative integral invariant* in the sense of Poincaré [7]:

$$\int \Omega = \int w_\beta\,\delta x^\beta. \tag{2.29}$$

Denoting by $d\Omega$ the exterior differential of the form Ω, we have in local coordinates:

$$d\Omega = dw_\beta \wedge dx^\beta = \frac{1}{2}\Big[\partial_\beta w_\alpha - \partial_\alpha w_\beta\Big]dx^\beta \wedge dx^\alpha. \tag{2.30}$$

To the form $d\Omega$ is associated the antisymmetric tensor of components:

$$\omega_{\beta\alpha} = \partial_\beta w_\alpha - \partial_\alpha w_\beta. \tag{2.31}$$

It is easy to verify that these components coincide with the vorticity tensor components defined by (2.24). Unlike the vorticity tensor $\omega_{\beta\alpha}$, the vorticity vector ω^β does not remains the same upon the conformal transformations (2.27)–(2.28).

2.3.3 Application to the Gödel model

On the modified Gödel manifold (M,$\boldsymbol{g}_\mathrm{G}$), the components of the unit 4-vector $\boldsymbol{w}_\mathrm{G}$ tangent to world lines of matter (1.3) (1.4) are here:

$$(w^\mu)_\mathrm{G} = \mathrm{e}^U(u^\mu)_\mathrm{G} = \mathrm{e}^U(1,0,0,0), \tag{2.32}$$

$$\left(w_\beta\right)_\mathrm{G} = \mathrm{e}^U\left(u_\beta\right)_\mathrm{G} = \mathrm{e}^{-U}(1,0,\mathrm{e}^{x_1},0). \tag{2.33}$$

Notice that the contravariant components $(u^\mu)_\mathrm{G}$ are all constant.

In this *particular case*, according to (2.23), one has

$$(\mathring{u}_\alpha)_\mathrm{G} = \partial_\alpha U = 0, \quad \text{i.e., } U \text{ is constant.}$$

By concatenation, the conformal factor $\exp U$ reduces to a constant and coincides with Gödel's choice $a = const$.

So the vorticity magnitude of the fluid's matter remains as in the initial theory:

$$\omega_\mathrm{G} = \left(g_{\alpha\mu}\omega^\alpha\omega^\mu\right)^{1/2}_\mathrm{G} = \frac{a}{\sqrt{2}}. \tag{2.34}$$

On the other hand, we note that the covariant components of the velocity $(u_\beta)_\mathrm{G}$ are not all constant.

This means that the conformal geodesics principle holds within our theory.

In other words, we clearly see that Gödel's proposed solution is only a (very limited) special case (contravariant velocity components) which therefore reveals a patent lack of generality.

Therefore, Gödel's theory ought to be embedded in a broader scheme implying a conformal metric $\left(ds^2\right)'$ as we inferred above.

Note: one of the *Kretschmann scalar* is an invariant only for $\omega_{\mathrm{G}} : R_{\alpha\beta\gamma\delta}T^{\alpha\beta\gamma\delta} = 12\omega_{\mathrm{G}}{}^4$.

2.4 Chronal horizon

With Gödel one defines new (cylindrical) coordinates (t, r, ϕ, y) by setting:

$$\mathrm{e}^{x_1} = \cosh 2r + \cos\phi \sinh 2r, \tag{2.35}$$

$$x_2 \mathrm{e}^{x_1} = \sqrt{2}\sin\phi \sinh 2r, \tag{2.36}$$

$$\tan\frac{1}{2}\left[\phi + \frac{x - 2t}{\sqrt{2}}\right] = \mathrm{e}^{-2r}\tan\frac{\phi}{2}, \tag{2.37}$$

$$2z = x_3. \tag{2.38}$$

Within the framework of our theory, these coordinates lead to the line element:

$$ds^2 = 4\mathrm{e}^{2U}(x)\Big[dt^2 - dr^2 - dz^2 + \\ + \left(\sinh^4 r - \sinh^2 r\right)d\phi^2 + 2\sqrt{2}\sinh^2 r\, d\phi\, dt\Big]. \tag{2.39}$$

This metric still exhibits the rotational symmetry of the solution about the axis $r = 0$, since we clearly see that the components of the metric tensor do not depend on ϕ.

For $r \geqslant 0$, we have: $0 \leqslant \phi \leqslant 2\pi$. If a curve r_{G} is defined by: $\sinh^4 r = 1$, that is

$$r_{\mathrm{G}} = \ln\left(1 + \sqrt{2}\right), \tag{2.40}$$

then any curve $r > \ln(1 + \sqrt{2})$, i.e. $(\sinh^4 r - \sinh^2 r) > 0$ materialized in the "plane" $t = const.$ (or zero t), is a *closed timelike curve*.

The radius r_{G} referred to as the *Gödel radius* induces a *light-like curve* or *closed null curve*, where the light cones are tangential to the plane of constant (or zero) t.

The photons trajectories reaching this radius are closing up, therefore r_{G} constitutes a *chronal horizon* beyond which an observer located at the origin ($r = 0$) cannot detect them.

With increasing $r > r_{\mathrm{G}}$ the light cones continue to keel over and their opening angles widen until their future parts reach the negative values of t.

In this *achronal domain*, any trajectory is a closed time-like curve and s' is extended over a full cycle.

As a result, the integral U performed over the closed path has no endpoints and is thus expressed in the form:

$$U = \oint\left[\frac{dP}{\rho + P}\right] + const. \tag{2.41}$$

However, the pressure P which is fluctuating along the closed path remains at the same averaged value for the whole cycle and may be then regarded as globally constant.

In this case, the first term in the r.h.s. of (2.35) vanishes implies $U = const.$, and the conformal factor ($\exp U$) may coincide again with Gödel's choice $a = const.$

Therefore, for $r > r_{\mathrm{G}}$, the acceleration of flow lines of matter is always zero whatever the components of w_{G}. Because of this, all closed timelike curves can no longer be derived from the geodesic principle calculation developed above.

By introducing the pressure in the Gödel model, we clearly put in evidence the difference between the geodesics and the closed time-like curves.

This was mathematically outlined in [8] but no explanation was provided as *why* this difference arises.

Conclusion

When Gödel wrote down his metric he was led to introduce a distinctive constant factor a in order to re-transcript the field equations with a cosmological constant along with additional constraints.

Our theory is free of all these constraints and moreover it provides a physical meaning to the a term. Inspection shows that by substituting a conformal factor to the constant a induces the field equations with a pressure like term which was wrongly interpreted by Gödel as the cosmological constant of the universe.

In fact, he empirically assembled the pieces of the constant matter density and curvature scalar in order to conveniently cope with the field equations precisely written with the cosmological constant.

In contrast, the reconstructed Gödel metric is here a straightforward solution to these equations and as such it can be reproduced like any other metric without referring to any cosmological model whatsoever.

The metric still exhibits a rotation which allows for the existence of *close timelike curves (CTCs)* since the light cone opens up and tips over, as the Gödel's circular coordinate radius increases within the cylindrical coordinates representation.

It seems that the first model exhibiting this property was pioneered by the German mathematician C. Lanczos in 1924 [9], and later rediscovered in a new form by the dutch physicist W. J. Van Stockum in 1937 [10].

However, the existence of *CTCs* satisfying the Einstein's equations remained so far a stumbling block for most of physicists because it should imply the possibility to travel back and forth in time.

The time travel possibility, was quoted as a pure mathematical "exercise" unrealistic in nature because it was deemed to describe a hypothetical universe contradicting the standard model in expansion as we observe it. Moreover, defining an absolute time is not readily applicable in Gödel space-time.

In here, the cards are now somewhat reshuffled: the Gödel model does not describe any sort of universe and the relevant metric can be applied as any other metrics like for example the Schwarzschild, the Kerr or the Alcubierre's ones.

Under these circumstances, why not considering the Gödel model as a potential time machine?

A typical example of such possible time machine is given by the cylinder system elaborated by the American physicist F. J. Tipler in 1974 [11].

It describes an infinitely long massive cylinder spinning along its longitudinal axis which gives rise to the "frame dragging" effect. If the rotation rate is fast enough the light cones of objects in cylinder's vicinity become tilted. Tipler suggested that a finite cylinder might also produce *CTCs* which was objected by Hawking who argued that any finite region containing *CTCs* would require negative energy density produced by a so-called "exotic matter" which violates all *energy conditions* [12].

The same kind of negative energy is needed to sustain a coupled system of *Lorentzian wormholes* designed to create a time machine as suggested in [13].

In all cases, feasibility and related causality paradoxes seemed to have been killed once for good by Hawking through a specific vacuum fluctuations mechanism that impedes any attempt to travel in the past [14].

Several authors have however recently challenged if not rejected this statement [15, 16].

These constraints do not apply in the present theory.

For a thorough study covering *CTCs* questions one can refer to [17, 18].

Submitted on May 24, 2021

References

1. Gödel K. An example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation. *Review of Modern Physics*, 1949, v.21(3).
2. Marquet P. Gödel's Universe Revisited. *Progress in Physics*, 2014, v.10, 259–262.
3. Eisenhart L.P. *Trans. Americ. Math. Soc.*, 1924, v.26, 205–220.
4. Synge J.L. *Proc. London Math. Soc.*, 1937, v.43, 37–416.
5. Lichnerowicz A. Les Théories Relativistes de la Gravitation et de l'Electromagnétisme. Masson et Cie, Paris, 1955.
6. Kramer D., Stephani H., Herlt E., Mac Callum M. Exact Solutions of Einstein's Field Equations. Cambridge University Press, Cambridge, 1982.
7. Cartan E. Leçons sur les invariants intégraux. A. Hermann, Paris, 1922.
8. Chandrasekhar S., Wright J.P. The geodesics in Gödels's universe. *Proc. of the National Academy of Sciences*, 1961, v.47(3), 341–347.
9. Lanczos C. Über eine Stationäre Kosmolologie im Sinne der Einsteinischer Gravitationstheorie. *Zeitschrift für Physik*, 1924, Bd.21, 73.
10. Van Stockum W. J. The gravitational field of a distribution of particles rotating around an axis of symmetry. *Proc. Roy. Soc. Edinburgh*, 1937, v.A57, 135.
11. Tipler F.J. Rotating cylinders and the possibility of global causality violation. *Phys. Rev. D.*, 1974, 9(8), 2203–2206.
12. Marquet P. Exotic Matter: A New Perspective. *Progress in Physics*, 2017, v.13, 174–179.
13. Morris M. S., Thorne K. S., Yurtsever U. Wormholes, Time machine and the weak energy condition. *Physical Rev. Letters*, 1988, v.61(13), 1446–1449.
14. Hawking S. W. Chronology protection conjecture. *Phys. Rev. D.*, 1991, v.46(2), 603–610.
15. Kim Sung Won, Thorne K. S. Do vacuum fluctuations prevent the creation of closed time-like curves? *Physical Review D*, 1991, v.43(12), 3929–3947.
16. Li Li-Xin. Must time machine be unstable against vacuum fluctuations? 2006, arXiv: gr-qc/9703024.
17. Friedmann J., Morris M.S., Novikov I., Echeverria F., Klinkhammer G., Thorne K.S., Yurtsever U. Cauchy problem in space-times with closed timelike curves. *Physical Rev. D*, 1990, v.42(6), 1915.
18. Thorne K. S. Closed Timelike Curves. Caltech, 1993, GRP-340.
19. Hawking S.W., Ellis G.F.R. The Large Scale Structure of Space-Time. Cambridge University Press, Cambridge, 1987.

Appendix

In a *holonomic frame* defined on (M,$\boldsymbol{g}$), the unit vectors are normalized so that:

$$g_{\mu\nu} u^\mu u^\nu = g^{\mu\nu} u_\mu u_\nu = 1. \tag{A.1}$$

By differentiating we get:

$$u^\mu \nabla_\nu u_\mu = 0. \tag{A.2}$$

Let us consider the following tensor which describes a homogeneous perfect fluid with density ρ and with pressure P:

$$T_{\mu\nu} = (\rho + P) u_\mu u_\nu - P g_{\mu\nu}. \tag{A.3}$$

The conservation equations are written:

$$\nabla_\mu [(\rho + P) u^\mu u_\nu] = \nabla_\mu \Big(P \delta_\nu{}^\mu \Big). \tag{A.4}$$

Setting the vector K_ν such that

$$(\rho - P) K_\nu = \nabla_\mu \Big(P \delta_\nu{}^\mu \Big), \tag{A.5}$$

$$\nabla_\mu [(\rho - P) u^\mu u_\nu] = (\rho + P) K_\nu, \tag{A.6}$$

$$\nabla_\mu [(\rho + P) u^\mu] u_\nu + (\rho + P) u^\mu \nabla_\mu u_\nu = (\rho + P) K_\nu. \tag{A.7}$$

Multiplying through with u^ν, and taking into account (A.2), we obtain after dividing by $(\rho + P)$:

$$u^\mu \nabla_\mu u_\nu = \Big(g_{\mu\nu} - u_\mu u_\nu \Big) K^\mu = h_{\mu\nu} K^\mu. \tag{A.8}$$

The flowlines everywhere tangent to the vector u^μ are determined by the differential equations (A.8).

K^μ only depends on x^μ and since: $h_{\mu\nu} K^\mu = K_\nu = \partial_\nu \dfrac{P}{\rho - P}$, we set

$$K_\nu = \partial_\nu U, \tag{A.9}$$

with

$$U = \int \frac{dP}{\rho + P}. \tag{A.10}$$

When the fluid pressure is function of the density, the 4-vector $\partial_\nu U$ is regarded as the *4-acceleration vector* $\mathring{u}_\nu$ of the flow lines given by the pressure gradient orthogonal to those lines [19, p.70].

A Statistical Approach to Two-particle Bell Tests

Xianming Meng

Research School of Physics, Australian National University, Canberra, ACT2601. E-mail: xianming.meng@anu.edu.au

Extensive experimental tests of the Bell inequality have been conducted over time and the test results are viewed as a testimony to quantum mechanics. In considering the close tie between quantum mechanics and statistical theory, this paper identifies the mistake in previous statistical explanation and uses an elegant statistical approach to derive general formulas for two-particle Bell tests, without invoking any wavefunctions. The results show that, for the special case where the spins/polarizations are in the same, opposite, or perpendicular directions, the general formulas derived in this paper convert to quantum predictions, which are confirmed by numerous experiments. The paper also investigates the linkages between the statistical and quantum predictions and finds that vector decomposition and probability law are at the heart of both approaches. Based on this finding, the paper explains statistically why the local hidden variable theory fails the Bell tests. The paper has important implications for quantum computing, quantum theory in general, and the role of randomness and realism in physics.

1 Introduction

The extensive study on Bell tests originated from the 1935 paper by Einstein *et al* [1], which claimed that physical reality can be predicted with certainty and that the uncertain nature of quantum prediction is due to incomplete information or the act of local hidden variables. Bohm [2] proposed a thought experiment to test the local hidden variable (LHV) theory and quantum mechanism, but this thought experiment was impractical to implement. In 1964, John Bell [3] developed the Bell inequality from the LHV theory as a testing tool: if the inequality is violated, the LHV theory is disproved. In 1969 Clauser *et al* [4] extended the Bell inequality to an experimentally testable version. Freedman and Clauser [38], Aspect [5, 6] and many others used this version to test the inequality and convincingly rejected it. Numerous experiments on Bell tests [7–21] have been conducted to close the "loopholes" in testing. Since almost all testing results are consistent with the quantum mechanical prediction, they are viewed as a testimony to quantum mechanism.

It is well known that quantum mechanics has a close tie with probability theory. The author suspects that both quantum mechanics and statistics mechanics may essentially be the same in the case of the Bell tests, and therefore identified the mistakes in previous statistical explanation and derived a statistical prediction for two-particle Bell tests. It is revealed that the quantum prediction of the Bell test results is a special case of the statistical prediction. By comparing the statistical and quantum derivations, the author further demonstrates that the essence of quantum prediction is probability law, and that quantum entanglement in two-particle Bell tests is nothing mysterious but an alternative expression for statistical correlation (i.e. there is no difference between statistical and quantum correlations). When the correlated particles are separated and facing different conditions (e.g. polarizers of different orientations), probability law can still maintain their correlation.

The paper is organized as follows: Section 2 demonstrates the deterministic or uncorrelated nature of the Bell inequality and reveals the mistakes in the previous statistical approach. Based on a general case of spin or polarization, Section 3 derives a statistical prediction for Bell tests for all possible uncorrelated and correlated particle pairs. Section 4 explores the linkage between the quantum and statistical predictions, while Section 5 uses the statistical approach to explain the results of representative two-particle Bell tests. Section 6 concludes the paper.

2 Deterministic or uncorrelated nature of the Bell inequality

Realism and localism play a key role in deriving the Bell inequality. The usual assumption for derivation is that at location A, a setting a (e.g. the direction of the spin/polarization analyser) leads to an experimental outcome $A(a)$, while setting b at location B leads to outcome $B(b)$, with the joint outcome being $E(a, b) = A(a)\,B(b)$. Since a setting leads to an outcome with certainty, the outcome is predetermined by the settings. This fits with the idea of determinism or realism. Moreover, the outcome at a location is determined only by the setting at that location, e.g. $A(a)$ is determined by local setting a at location A, not by setting b at location B. This is localism.

If settings a and b can be changed to a' and b', respectively, we can have joint outcomes:

$$E(a, b') = A(a)\,B(b')$$

$$E(a', b) = A(a')\,B(b)$$

$$E(a', b') = A(a')\,B(b')\,.$$

We further assume that the detected outcome at any setting is between -1 and +1, namely $|A| \leq 1$, $|B| \leq 1$. With these assumptions, we can have:

$$\begin{aligned}
&E(a,b)-E(a,b')=A(a)\,B(b)-A(a)\,B(b')\\
&=A(a)\,B(b)-A(a)\,B(b')+[A(a)\,B(b)][A(a')\,B(b')]\\
&-[A(a)\,B(b)][A(a')\,B(b')]
\end{aligned}$$

or

$$\begin{aligned}
E(a,b)-E(a,b') \quad &= A(a)\,B(b)\,[1+A(a')\,B(b')]-\\
&-A(a)\,B(b')\,[1+A(a')\,B(b)]\,.
\end{aligned} \tag{1}$$

In absolute value, we can write:

$$\begin{aligned}
\left|E(a,b)-E(a,b')\right| \quad &\le |A(a)\,B(b)| * |1+A(a')\,B(b')| \;+\\
&+|A(a)\,B(b')| * |1+A(a')\,B(b)|\;.
\end{aligned} \tag{2}$$

We have changed the negative sign on the right-hand side of (1) to a positive sign in (2) because $A(a)\,B(b')$ can be negative. Since the values of $A(a)$, $B(b)$, $A(a')$, and $B(b')$ are all between -1 and 1, we have:

$$|A(a)\,B(b)| \le 1 \quad \text{and} \quad \left|A(a)\,B(b')\right| \le 1\,.$$

As such, the inequality can be written as:

$$\begin{aligned}
\left|E(a,b)-E(a,b')\right| &\le \left|1+A(a')\,B(b')\right|+\left|1+A(a')\,B(b)\right|\\
&=2\pm\left|A(a')\,B(b')+A(a')\,B(b)\right|
\end{aligned}$$

or

$$\left|E(a,b)-E(a,b')\right| \le 2\pm\left|E(a',b')+E(a',b)\right|\,. \tag{3}$$

On the right-hand side of (3), we used the "±" sign because both $A(a')\,B(b')$ and $A(a')\,B(b)$ can be negative (leading to negative sign) or positive (leading to positive sign). There are two boundaries in the above inequality. If the lower boundary is satisfied, the inequality holds, so we have arrived at the Bell inequality:

$$\left|E(a,b)+E(a',b')+E(a'b)-E(a,b')\right| \le 2\,. \tag{4}$$

To incorporate a hidden variable into the inequality, most researchers introduced a random variable. For example, Bell [3,22] and Clauser *et al* [4] added to the experiments a hidden variable λ, which has a normalized probability distribution:

$$\int_{-\infty}^{\infty} p(\lambda)\,d\lambda = 1\,.$$

With the added hidden variable, Bell [3,22] expressed the expected values of coincidence at the different settings a, a', b and b' as follows:

$$E(a,b)=\int_{-\infty}^{\infty} A(a,\lambda)\,B(b,\lambda)\,p(\lambda)\,d\lambda \tag{5}$$

$$E(a,b')=\int_{-\infty}^{\infty} A(a,\lambda)\,B(b',\lambda)\,p(\lambda)\,d\lambda \tag{6}$$

$$E(a',b)=\int_{-\infty}^{\infty} A(a',\lambda)\,B(b,\lambda)\,p(\lambda)\,d\lambda \tag{7}$$

$$E(a',b')=\int_{-\infty}^{\infty} A(a',\lambda)\,B(b',\lambda)\,p(\lambda)\,d\lambda\,. \tag{8}$$

Using the same procedure that was used to derive the Bell inequality for (3) – the deterministic case, Bell ([22, pp. 178–179]) derived (the notations are slightly changed for contemporary readers):

$$\begin{aligned}
&E(a,b)-E(a,b') \;=\\
&=\textstyle\int_{-\infty}^{\infty} A(a,\lambda)\,B(b,\lambda)\,p(\lambda)\,d\lambda\;-\\
&-\textstyle\int_{-\infty}^{\infty} A(a,\lambda)\,B(b',\lambda)\,p(\lambda)\,d\lambda\\
&=\textstyle\int_{-\infty}^{\infty}[A(a,\lambda)\,B(b,\lambda)-A(a,\lambda)\,B(b',\lambda)\;+\\
&+A(a,\lambda)\,B(b,\lambda)\,A(a',\lambda)\,B(b',\lambda)\;-\\
&-A(a,\lambda)\,B(b,\lambda)\,A(a',\lambda)\,B(b',\lambda)]\,p(\lambda)\,d\lambda\\
&=\textstyle\int_{-\infty}^{\infty} A(a,\lambda)\,B(b,\lambda)\,[1+A(a',\lambda)\,B(b',\lambda)]\,p(\lambda)\,d\lambda\;-\\
&-\textstyle\int_{-\infty}^{\infty} A(a,\lambda)\,B(b',\lambda)\,[1+A(a',\lambda)\,B(b,\lambda)]\,p(\lambda)\,d\lambda\,.
\end{aligned} \tag{9}$$

In terms of absolute value, we have:

$$\begin{aligned}
&|E(a,b)-E(a,b')|\\
&\le\left|\int_{-\infty}^{\infty} A(a,\lambda)\,B(b,\lambda)\,[1+A(a',\lambda)\,B(b',\lambda)]\,p(\lambda)\,d\lambda\right|\\
&+\left|\int_{-\infty}^{\infty} A(a,\lambda)\,B(b',\lambda)\,[1+A(a',\lambda)\,B(b,\lambda)]\,p(\lambda)\,d\lambda\right|\\
&\le\left|\int_{-\infty}^{\infty}[1+A(a',\lambda)\,B(b',\lambda)]\,p(\lambda)\,d\lambda\right|\\
&+\left|\int_{-\infty}^{\infty}[1+A(a',\lambda)\,B(b,\lambda)]\,p(\lambda)\,d\lambda\right|\\
&=2\pm|E(a',b')+E(a',b)|\,.
\end{aligned}$$

Rearranging the above inequality as before, we can obtain the same inequality as (4).

From the above derivation, one may notice that the same term $\int_{-\infty}^{\infty} p(\lambda)\,d\lambda$ is added to outcomes of the different settings and then this term is filtered out in the end by the definition of expected values in (7) and (8). As such, the added hidden variable and probability are only additional statistical noise, which does not change the deterministic nature of the resulting inequality.

Later, Bell and others [28–30] moved on to a version of the Bell inequality based on joint and conditional probabilities. However, they used the same assumption that the distribution of hidden variable λ is UNRELATED to local settings. This assumption apparently contradicts the concept of a local variable. Ironically, the assumption is often regarded as a feature of a local variable. Myrvold *et al* [23] used a different approach. Instead of concerning the probability distributions of λ conditioned on settings, they conditioned the experimental outcomes on hidden variable λ. Since they assigned no statistical property to λ, its behaviour is unknown, so its role in their derivation is negligible, or not essential at least.

To present a genuine statistical event, one should allow the probability density λ to vary with the local settings. In other words, the probability of value λ must be conditioned on the settings, i.e. for settings a and b, we have the probability $p(\lambda\,|\,a)$ and $p(\lambda\,|\,b)$, respectively. The probability of the joint outcome of settings a and b should be $p(\lambda\,|\,a,b)$. Similarly, we have $p(\lambda\,|\,a,b')$, $p(\lambda\,|\,a',b)$, $p(\lambda\,|\,a',b')$ for other joint settings. As such, the expected joint detection should be:

$$E(a,b) = \int_{-\infty}^{\infty} A(a,\lambda)\, B(b,\lambda)\, p(\lambda \mid a,b)\, d\lambda$$

$$E(a,b') = \int_{-\infty}^{\infty} A(a,\lambda)\, B(b',\lambda)\, p(\lambda \mid a,b')\, d\lambda$$

$$E(a',b) = \int_{-\infty}^{\infty} A(a',\lambda)\, B(b,\lambda)\, p(\lambda \mid a',b)\, d\lambda$$

$$E(a',b') = \int_{-\infty}^{\infty} A(a',\lambda)\, B(b',\lambda)\, p(\lambda \mid a',b')\, d\lambda\,.$$

Using this new definition of expected values, the terms for the probability of λ are different for each joint setting and thus cannot be filtered out. As a result, the Bell inequality cannot be derived.

However, one may further assume that the joint probability of outcome at joint setting a and b is the multiplication of probabilities of outcomes at each setting, namely:

$$p(\lambda\,|\,a,b) = p(\lambda\,|\,a)\, p(\lambda\,|\,b) \tag{10}$$

where $0 \leq p(\lambda\,|\,a) < 1$; $0 \leq p(\lambda\,|\,b) < 1$; $\int_{-\infty}^{\infty} p(\lambda\,|\,a)\, d\lambda = 1$; and $\int_{-\infty}^{\infty} p(\lambda\,|\,b)\, d\lambda = 1$.

Applying the same method for joint settings a and b', a' and b, and a' and b', we have:

$$p(\lambda\,|\,a,b') = p(\lambda\,|\,a)\, p(\lambda\,|\,b')$$

$$p(\lambda\,|\,a',b) = p(\lambda\,|\,a')\, p(\lambda\,|\,b)$$

$$p(\lambda\,|\,a',b') = p(\lambda\,|\,a')\, p(\lambda\,|\,b')\,.$$

Based on the above joint probabilities, we can calculate $E(a,b)$, $E(a,b')$, $E(a',b)$ and $E(a',b')$. Following the same procedure as in deriving (9), we can derive the Bell inequality (4).

As we see, (10) is crucial for deriving the Bell inequality from a statistical point of view. However, the expression of joint probability as a product of the probability of outcome of two experiments is not without a condition. The well-known but often neglected condition is that the two experiments involved in the joint probability calculation in (10) must be totally unrelated, i.e. independent random experiments. Applying this condition to the Bell tests, the requirement is that the probabilities of outcomes at different locations/settings are independent of each other, so "local" means "uncorrelated". This interpretation gives the alternative condition for the Bell inequality. That is, if the outcomes are not deterministic, the outcomes at two different settings should not be correlated.

The common wisdom is that, during a Bell test, the experiments at different locations A and B are apparently independent because the orientations of the polarizers at A and B are changed independently and randomly. However, the independence of settings are not the full condition for independent experiments because local settings are only one element of the polarization experiments. The other element is the light source. In fact, correlated source particles are used in all Bell tests conducted so far, so the experiments conducted at different locations are not independent. Since the experiments based on different settings are correlated by source particles, the joint probability in a Bell test should be calculated based on conditional probability:

$$p_{a,b} = p_a * p_{b|a}$$

or

$$p_{a,b} = p_b * p_{a|b}\,.$$

Similar mistakes are also commonly made in treating the expected value of joint events as being the multiplication of the expected values of separate events. Due to the statistical nature of the polarization experiments, we need to allow one setting to generate different results, e.g. experiments based on setting a can have results $A_1(a), A_2(a), \ldots A_n(a)$, so the expected value for results of setting a can be expressed as:

$$E(a) = \frac{1}{n} \sum_i A_i(a)\,. \tag{11}$$

We can also write the expected value for results of setting b as:

$$E(b) = \frac{1}{n} \sum_i B_i(b)\,. \tag{12}$$

Indeed, Bell [22, p. 178] realized the importance of introducing (11) and (12) for $E(a)$ and $E(b)$. However, with no precondition being specified, he assumed the following equality as the base for deriving the Bell inequality:

$$E(a,b) = E(a) * E(b)\,. \tag{13}$$

The above equation is used by numerous researchers on Bell tests, but the equation is not unconditional. Statistically, we can expand the expected values as:

$$E(a,b) = \frac{1}{n} \sum_i A_i(a)\, B_i(b) \tag{14}$$

$$E(a) * E(b) = \frac{1}{n^2} \sum_i A_i(a) \sum_i B_i(b)\,. \tag{15}$$

Apparently, $E(a,b) \neq E(a) * E(b)$ in general cases. A special statistical case where $E(a,b) = E(a) * E(b)$ holds is when the outcomes of $A_i(a)$ are independent of (or not correlated to) the outcomes of $B_i(b)$. In this special case, the Bell inequality will hold. If $E(a)$ and $E(b)$ are correlated, we must use the conditional expected values that reflect the correlations between two experiments.

From the above discussion, we can conclude that the Bell inequality does not allow for a probabilistic nature (or correlation, to be exact) because it is based on determinism or realism. To allow for the Bell inequality in a statistical experiment, one must satisfy the condition for (13), which in turn requires that there is no correlation between $A_i(a)$ and $B_i(b)$. In terms of quantum mechanics terminology, if particles 1 and 2 are in separable (uncorrelated) states, the Bell inequality is valid, otherwise (if particles 1 and 2 are in entangled states), the Bell inequality will be violated.

3 A statistical interpretation of spin/polarization correlation

A statistical presentation of Bell tests seems to be complicated because it involves many random settings, such as random directions of polarizers and random polarization of light or spins of particles. Moreover, spins and polarizations have different features. After trying a number of methods, the author has arrived at a remarkably simple and elegant approach for deriving the statistical prediction.

The difference between polarization and spin is that spins in opposite directions have different values while polarizations in the opposite directions are viewed as being the same. In other words, the spin direction in a plane can have a 360° variation while the polarization direction varies only within 180°, so the case of polarization is a reduced case of spin. For generality, this section focuses on deriving the results for the case of spin, and then shows how the results can be applied to the case of polarization.

There are various types of spin analyzer/detector [24–27], but all spin detectors rely on a differing scattering cross section for spin polarized particles. During spin detection, the travel direction of the particle and the detector orientation form a plane, in which the particles are reflected and detected [25]. The spin polarized particles will cause asymmetric reflection, and the asymmetric results indicate the detected spin direction. Essentially, a spin analyser works similarly to a polarizer for light, but the analyser can identify the spin direction along the given detection orientation. Consequently, we use a polarizer with an arrow (a vector) to represent a spin analyzer.

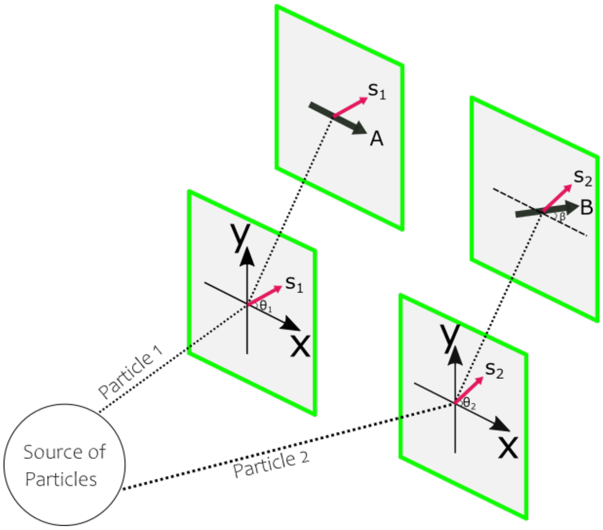


Fig. 1: Measuring spin directions

Fig. 1 shows a general case where the particles of the different spin directions are measured by the two spin analyzers in a Bell test experiment. Two spins, s_1 and s_2, and two spin analyzers, A and B, are positioned in different directions. The spin directions of particles 1 and 2 form an angle of θ_1 and θ_2, respectively, with the x-axis. For simplicity, we assume that s_1 and s_2 are unit vectors, and that spin analyzer A is placed in the direction of the x-axis while spin analyzer B forms an angle of β with the x-axis. Given this setting, the component of s_1 detected by A is $E(A) = \cos\theta_1$. Similarly, the angle between s_2 and the spin analyzer B is $\theta_2 - \beta$, so the component of s_2 detected by B is $E(B) = \cos(\theta_2 - \beta)$.

There are two types of correlation measurement in the Bell tests. One is the joint detection counts normalized on the separate detection counts at each setting. The other is the joint detection rate normalized on the emission rate at the particle source. We address them in turn.

3.1 Correlation normalized on outcomes at each setting

This measurement fits with the standard definition of correlation, so we can calculate the expected value, variance and covariance and then obtain correlation. Since the source emits particles of random spin directions, the expected values and variances can be obtained by integrating $E(A)$ and $E(B)$ over the spin angles θ_1 and θ_2 in the range of $0 - 2\pi$ for particles 1 and 2.

$$\langle E(A)\rangle = \frac{\int_0^{2\pi} E(A)\,d\theta_1}{\int_0^{2\pi} d\theta_1} = \frac{\int_0^{2\pi} \cos\theta_1\,d\theta_1}{\int_0^{2\pi} d\theta_1} = \left.\frac{\sin\theta_1}{\theta_1}\right|_0^{2\pi} = 0$$

$$\begin{aligned}\mathrm{var}(A) &= \frac{\int_0^{2\pi} [\cos\theta_1 - \langle E(A)\rangle]^2\,d\theta_1}{\int_0^{2\pi} d\theta_1}\\ &= \frac{1}{2\pi}\int_0^{2\pi} \cos^2\theta_1\,d\theta_1\\ &= \frac{1}{2\pi}\int_0^{2\pi} 0.5\,(\cos 2\theta_1 + 1)\,d\theta_1 = 0.5\end{aligned}$$

$$\langle E(B)\rangle = \frac{\int_0^{2\pi} E(B)\,d\theta_2}{\int_0^{2\pi} d\theta_2} = \frac{\int_0^{2\pi} \cos(\theta_2 - \beta)\,d\theta_2}{\int_0^{2\pi} d\theta_2} = 0$$

$$\begin{aligned}\mathrm{var}(B) &= \frac{\int_0^{2\pi}[\cos(\theta_2-\beta)-\langle E(B)\rangle]^2\,d\theta_2}{\int_0^{2\pi} d\theta_2}\\ &= \frac{1}{2\pi}\int_0^{2\pi}\cos^2(\theta_2-\beta)\,d\theta_2\\ &= \frac{1}{2\pi}\int_0^{2\pi}0.5\,[\cos 2(\theta_2-\beta)+1]\,d\theta_2 = 0.5\,.\end{aligned}$$

If the two particles are uncorrelated, θ_1 and θ_2 can vary independently, so the covariance can be calculated through a double integral:

$$\begin{aligned}&\mathrm{cov}(A,B) =\\ &= \frac{\iint_0^{2\pi}[\cos\theta_1-\langle E(A)\rangle][\cos(\theta_2-\beta)-\langle E(B)\rangle]\,d\theta_1\,d\theta_2}{\iint_0^{2\pi} d\theta_1\,d\theta_2}\\ &= \frac{1}{(2\pi)^2}\int_0^{2\pi}\cos\theta_1\,d\theta_1\int_0^{2\pi}\cos(\theta_2-\beta)\,d\theta_2 = 0\,.\end{aligned}$$

The zero covariance is expected because of the uncorrelated nature of s_1 and s_2 — the positive and negative joint detection counts will be largely cancelled out. If the two spins are correlated, θ_1 and θ_2 can still change randomly, but these two angles must keep the same difference, i.e. $\theta_2 = \theta_1 + \theta_0$, where θ_0 is the fixed relative angle between two spin directions. In this case, the covariance can be calculated by an integration over θ_1 (or θ_2):

$$\begin{aligned}&\mathrm{cov}(A,B) =\\ &= \frac{\int_0^{2\pi}[\cos\theta_1-\langle E(A)\rangle][\cos(\theta_1+\theta_0-\beta)-\langle E(B)\rangle]\,d\theta_1}{\int_0^{2\pi} d\theta_1}\\ &= \frac{1}{2\pi}\int_0^{2\pi}0.5\,[\cos(2\theta_1+\theta_0-\beta)+\cos(\beta-\theta_0)]\,d\theta_1\\ &= 0.5\cos(\beta-\theta_0)\,.\end{aligned}$$

As such, we have the following spin correlation:

$$\begin{aligned}E(A,B) &= \frac{\mathrm{cov}(A,B)}{[\mathrm{var}(A)]^{1/2}[\mathrm{var}(B)]^{1/2}}\\ &= \frac{0.5\cos(\beta-\theta_0)}{0.5^{0.5}*0.5^{0.5}} = \cos(\beta-\theta_0)\,. \qquad (16)\end{aligned}$$

Eq. (16) is a general result for joint detection for any given orientations of spin detectors. The application of this equation for special occasions can produce quantum predictions. For example, if two particles have the same spin, i.e. entangled particles of the same phase, we have $\theta_0 = 0$, $E(A,B) = \cos\beta$. If two particles have the opposite spin, i.e., negatively correlated particles, we have $\theta_0 = \pi$, $E(A,B) = -\cos\beta$. If the two spin vectors are perpendicular, $\theta_0 = \pi/2$, $E(A,B) = \cos(\pi/2-\beta) = \sin\beta$.

It is worth mentioning that some researchers used light intensity correlation instead of the expected-value correlation for polarization Bell test. For example, Ou and Mandel [31] and Rarity and Tapster [35] regarded the joint detection probability of photons as being proportional to the intensity correlation of light. This approach is misplaced. For polarization experiments, one or more photons (assuming perfect detection for the simplicity of an argument) pass through the polarizer, a positive detection will be recorded, so the intensity is not an appropriate measurement. One may argue that intensity is the square of amplitude so intensity can be used as the proxy of probability of photons passing through the polarizer, based on which the joint probability can be calculated. However, as explained in Section 2, the joint probability cannot be calculated through the multiplication of probabilities of separate detections because of the correlated particles in a Bell test. Since probability measures the average of the squared detection values, the intensity (or probability) correlation approach will produce totally different result from that in this paper. This can be shown in the following expression:

$$p_{AB} = p_A\,p_B = \langle E(A)^2\rangle\,\langle E(B)^2\rangle \neq \langle E(A)\,E(B)\rangle^2 = E(A,B)^2\,.$$

3.2 Correlation normalized on emissions at the source

For a Bell test, one needs to measure many pairs of particles of different spin directions with varied detector orientations. In this case, the joint detection rate is generally normalized on the emission rate at the source and the correlation is calculated based on the fixed axes.

Referring to Fig. 1, if the correlation is calculated based on x and y axes, the component detected by analyzer A and B needs to be further decomposed on the x-axis and y-axis:

$$E_{Ax} = E(A) = \cos\theta_1 \text{ and } E_{Ay} = 0$$

$$E_{Bx} = E(B)\cos\beta = \cos(\theta_2-\beta)\cos\beta$$

$$E_{By} = E(B)\sin\beta = \cos(\theta_2-\beta)\sin\beta\,.$$

Since no component on the y-axis is detected by analyzer A, the correlation (joint detection) on the y-axis is zero. On the other hand, both analyzers detect values on the x-axis, so the joint detection value is:

$$E_{AB} = E_{Ax}\,E_{Bx} = \cos\theta_1\cos(\theta_2-\beta)\cos\beta\,.$$

Since the correlation is based on the emissions at source, which are 100% detected (assuming all particles come to and are detected by either detector A or B), the variances are one and thus the correlation is equivalent to co-variance. If particles 1 and 2 are uncorrelated, the joint detection rate will be the value of E_{AB} integrated over both θ_1 and θ_2:

$$p_{AB} = \frac{\iint_0^{2\pi} E_{AB}\,d\theta_1\,d\theta_2}{\iint_0^{2\pi} d\theta_1\,d\theta_2}$$

$$
\begin{aligned}
&= \frac{\iint_0^{2\pi} \cos\theta_1 \cos(\theta_2 - \beta)\cos\beta \, d\theta_1 \, d\theta_2}{\iint_0^{2\pi} d\theta_1 \, d\theta_2} \\
&= \frac{\cos\beta}{(2\pi)^2} \int_0^{2\pi} \cos\theta_1 \, d\theta_1 \int_0^{2\pi} \cos(\theta_2 - \beta) \, d\theta_2 = 0 \, .
\end{aligned}
$$

The above result indicates that for uncorrelated particles, the joint detection rate is zero. This makes sense. Due to the uncorrelated random nature, the different detection counts will be washed out by the independent random changes in θ_1 and θ_2.

If two particles are correlated, i.e. $\theta_2 = \theta_1 + \theta_0$, we can obtain correlation by integrating E_{AB} over θ_1 (or θ_2) in range $0 - 2\pi$:

$$
\begin{aligned}
p_{AB} &= \frac{\int_0^{2\pi} E_{AB} \, d\theta_1}{\int_0^{2\pi} d\theta_1} \\
&= \frac{\int_0^{2\pi} \cos\theta_1 \cos(\theta_1 + \theta_0 - \beta)\cos\beta \, d\theta_1}{\int_0^{2\pi} d\theta_1} \\
&= \frac{\cos\beta}{2\pi} \int_0^{2\pi} 0.5 \left[\cos(2\theta_1 + \theta_0 - \beta) + \cos(\beta - \theta_0)\right] d\theta_1 \\
&= 0.5 \, \cos(\beta - \theta_0) \cos\beta \, .
\end{aligned}
\tag{17}
$$

The above result shows that when the two spin vectors are correlated, i.e., the value of θ_0 is fixed, the joint detection rate is determined only by correlation phase θ_0 and the angle β formed by the orientations of two spin detectors.

Eqs. (16) and (17) can also be applied to light polarization experiments. In the case of polarized light, it is tricky to derive the joint detection because the detected values have to be non-negative and thus are not consistent with the cosine functions for $E(A)$ and $E(B)$. The common approach (e.g. Aspect *et al* [4, 5]) is to define the result of no-detection as -1, instead of 0. In other words, when the light polarization is perpendicular to the orientation of detector, most likely no photon will be detected and thus a result of -1 with a 90° will be recorded. With this definition, all angles in (16) and (17) should be halved, and then the equation is equally applicable to the Bell tests with polarized light.

Where the two spin vectors are in the same directions (i.e. $\theta_0 = 0$), (17) becomes:

$$
p_{AB} = 0.5 \cos^2\beta = 0.25 \, (\cos 2\beta + 1) \, . \tag{18}
$$

In this special case, the joint detection rate can also be derived without integration, as shown in Fig. 2.

To present three random directions (i.e. the same direction of spin of the two particles, and the directions of the two spin analyzers A and B), we can fix one of them because only the relative angles between them matter. For convenience of presentation, we assume the spin vector $\overrightarrow{OV}$ to be a unit vector pointing to $V(a_x/\sqrt{2}, a_y/\sqrt{2})$, where a_x and a_y are unit vectors at x and y directions, respectively.

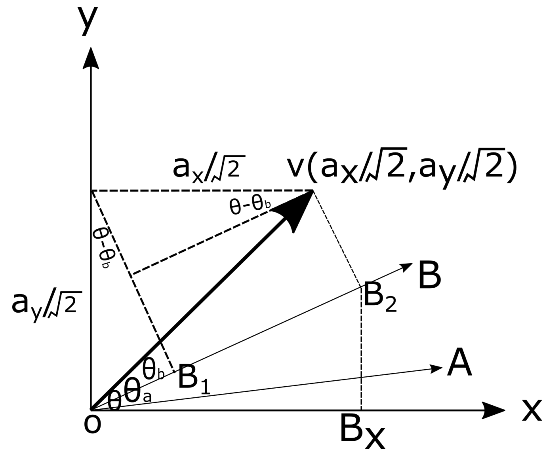


Fig. 2: Measuring the correlation of a particle pair of the same spin

The projection of the spin vector $\overrightarrow{OV}$ onto the B axis in Fig. 2 is:

$$
\overrightarrow{OB_2} = \overrightarrow{B_1B_2} + \overrightarrow{OB_1} = \left[\overrightarrow{a_x}\cos(\theta - \theta_b) + \overrightarrow{a_y}\sin(\theta - \theta_b)\right] / \sqrt{2} \, .
$$

This projection can be further projected onto the x-axis and y-axis and thus decomposed to two components $\overrightarrow{OB_x}$ and $\overrightarrow{OB_y}$, respectively ($\overrightarrow{OB_y}$ is not shown in Fig. 2 so as not to complicate the graph):

$$
\overrightarrow{OB_x} = \cos(\theta - \theta_b) \left[\overrightarrow{a_x}\cos(\theta - \theta_b) + \overrightarrow{a_y}\sin(\theta - \theta_b)\right] / \sqrt{2} \tag{19}
$$

$$
\overrightarrow{OB_y} = \sin(\theta - \theta_b) \left[\overrightarrow{a_x}\cos(\theta - \theta_b) + \overrightarrow{a_y}\sin(\theta - \theta_b)\right] / \sqrt{2} \, . \tag{20}
$$

Similarly, the projection of $\overrightarrow{OV}$ onto the A-axis can be decomposed into the x and y components of $\overrightarrow{OA_x}$ and $\overrightarrow{OA_y}$ respectively (not shown in Fig. 2):

$$
\overrightarrow{OA_x} = \cos(\theta - \theta_a) \left[\overrightarrow{a_x}\cos(\theta - \theta_a) + \overrightarrow{a_y}\sin(\theta - \theta_a)\right] / \sqrt{2}
$$

$$
\overrightarrow{OA_y} = \sin(\theta - \theta_a) \left[\overrightarrow{a_x}\cos(\theta - \theta_a) + \overrightarrow{a_y}\sin(\theta - \theta_a)\right] / \sqrt{2} \, .
$$

As such, the joint detection rate can be calculated as:

$$
\begin{aligned}
p_{AB} &= \overrightarrow{OA_x}\overrightarrow{OB_x} + \overrightarrow{OA_y}\overrightarrow{OB_y} \\
&= \cos(\theta - \theta_b) \left[\overrightarrow{a_x}\cos(\theta - \theta_b) + \overrightarrow{a_y}\sin(\theta - \theta_b)\right] / \sqrt{2} \\
&\times \cos(\theta - \theta_a) \left[\overrightarrow{a_x}\cos(\theta - \theta_a) + \overrightarrow{a_y}\sin(\theta - \theta_a)\right] / \sqrt{2} \\
&+ \sin(\theta - \theta_b) \left[\overrightarrow{a_x}\cos(\theta - \theta_b) + \overrightarrow{a_y}\sin(\theta - \theta_b)\right] / \sqrt{2} \\
&\times \sin(\theta - \theta_a) \left[\overrightarrow{a_x}\cos(\theta - \theta_a) + \overrightarrow{a_y}\sin(\theta - \theta_a)\right] / \sqrt{2} \\
&= 0.5 \, \left[\overrightarrow{a_x}\cos(\theta - \theta_b) + \overrightarrow{a_y}\sin(\theta - \theta_b)\right] \\
&\times \left[\overrightarrow{a_x}\cos(\theta - \theta_a) + \overrightarrow{a_y}\sin(\theta - \theta_a)\right] \cos(\theta_a - \theta_b) \\
&= 0.5 \cos^2(\theta_a - \theta_b)
\end{aligned}
$$

or

$$p_{AB} = 0.25\,[\cos 2(\theta_a - \theta_b) + 1]\,. \qquad (21)$$

Noting that $(\theta_a - \theta_b)$ is the angle formed by the orientations of two detectors A and B, we find that the above result is the same as (18). This joint probability of detection is exactly the same as the coincidence rate derived from quantum mechanics. The experiment by Aspect [5] confirmed this result.

The correlation function (16) and the joint detection rate (17) derived in this section are general results that are applicable to both uncorrelated or correlated polarization/spin of any phase differences. The results can be tested experimentally using the current Bell test techniques. The only change needed is to add a randomly controlled source polarizer for each of the two beams after the collimation lenses, but before the traditional Bell test polarizers. If the pair of source polarizers are randomly and separately controlled, i.e. their relative angle of polarization θ_0 varies randomly, the source particles are uncorrelated, so the joint detection rate will be zero for a large sample size. If the pair of source polarizers are controlled randomly but jointly, i.e. the relative polarization angle of the pair is fixed at any given value, the joint detection rate should be determined by the relative angle (θ_0) of the first pair of (source) polarizers and that (β) of the second pair, with the quantitative relations determined by (16) and (17).

4 Linkage between the statistical approach and quantum mechanics

From the previous section, we see that the simple statistical approach gives equivalent but more general results when they are compared with the predictions from quantum mechanics (QM). This is not a coincidence. This section shows that the statistical approach is at the heart of quantum mechanical prediction on Bell tests.

QM uses wavefunctions to represent the different states. For example, a wavefunction of a spin-up (or +1) state can be written in Dirac notation as $|0\rangle$, while spin-down (or -1) can be written as $|1\rangle$. The spin states can be projected to (or measured on) different axes and may result in different results. If Alice measures a spin state of $|0\rangle$ on the A-axis while Bob measures $|1\rangle$ on the B-axis, we can express this spin state as $|0\rangle \otimes |1\rangle$, or simply $|01\rangle$. A wavefunction $|01\rangle + |10\rangle$ indicates that the measurement on the A-axis is always opposite to the measured results on the B-axis, i.e. the measured results are negatively correlated. Similarly, the states in wavefunction $|00\rangle + |11\rangle$ are positively correlated. The states in this type of wavefunctions are called entangled states. On the other hand, a wavefunction of $|01\rangle + |00\rangle$ shows that while Alice's measurement is always $|0\rangle$, Bob's measurement can be either $|0\rangle$ or $|1\rangle$, so there is no correlation between the two measurement results. The states in this wavefunction are called separable states. In short, the entangled states are the QM expression for correlation.

Now we consider a normalized wavefunction of the positively entangled states: $\psi = (|00\rangle + |11\rangle)/\sqrt{2}$. If the states are measured by Alice on the A or x axis (both axes coincide, shown in Fig. 3), the possible outcome will be $\langle 0|\sigma^A|0\rangle = +1$ or $\langle 1|\sigma^A|1\rangle = -1$. Similarly, if the state is measured by Bob on the B-axis, the possible outcome will be $\langle 0|\sigma^B|0\rangle = +1$ or $\langle 1|\sigma^B|1\rangle = -1$. Since this is a wavefunction of positively entangled states, Alice and Bob will always obtain the same (positive or negative) measurement outcome. Bob's measurement can be decomposed to two components on the x-axis and y-axis: $\sigma^B = \sigma^B_x \cos\beta + \sigma^B_y \sin\beta$. Alternatively, we can write: $\langle 0|\sigma^B_x|0\rangle = \cos\beta$, $\langle 1|\sigma^B_x|1\rangle = -\cos\beta$, $\langle 0|\sigma^B_y|0\rangle = \sin\beta$, $\langle 1|\sigma^B_y|1\rangle = -\sin\beta$. Since Alice's measurement is on the x-axis, we have $\sigma^A = \sigma^A_x$.

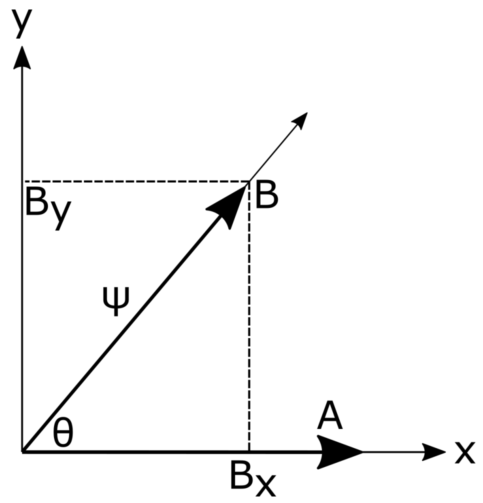


Fig. 3: Spin measurement for positively entangled particles

The correlation between the measurements of Alice and Bob can be calculated by the expected value of joint measurements: $\langle\sigma^A\sigma^B\rangle$. The QM calculation result is as follows:

$$\begin{aligned}
&\langle\sigma^A\sigma^B\rangle = \langle\psi|\sigma^A \otimes \sigma^B|\psi\rangle\\
&= 0.5\,(\langle 00| + \langle 11|)\sigma^A \otimes \sigma^B(|00\rangle + |11\rangle)\\
&= 0.5\,(\langle 00|\sigma^A \otimes \sigma^B|00\rangle + \langle 11|\sigma^A \otimes \sigma^B|00\rangle)\\
&+ \langle 00|\sigma^A \otimes \sigma^B|11\rangle + \langle 11|\sigma^A \otimes \sigma^B|11\rangle)\\
&= 0.5\,(\langle 0|\sigma^A|0\rangle\langle 0|\sigma^B|0\rangle + \langle 1|\sigma^A|0\rangle\langle 1|\sigma^B|0\rangle)\\
&+ \langle 0|\sigma^A|1\rangle\langle 0|\sigma^B|1\rangle + \langle 1|\sigma^A|1\rangle\langle 1|\sigma^B|1\rangle)\\
&= 0.5\,(\langle 0|\sigma^A|0\rangle\langle 0|\sigma^B|0\rangle + \langle 1|\sigma^A|1\rangle\langle 1|\sigma^B|1\rangle)\\
&= 0.5\,(\langle 0|\sigma^A_x|0\rangle\langle 0|\sigma^B_x|0\rangle + \langle 1|\sigma^A_x|1\rangle\langle 1|\sigma^B_x|1\rangle) = \cos\beta\,.
\end{aligned}$$

The above result is exactly the same as (16) with $\theta_0 = 0$, which was obtained from the much simpler statistical approach. A number of statistical features in the QM approach contribute to this same result. First, the calculation of the expected value in QM (i.e. $\langle\sigma^A\sigma^B\rangle = \langle\psi|\sigma^A \otimes \sigma^B|\psi\rangle$) is based on a probability-weighted average. Second, the rule of tensor product ($\langle 11 \,|\, \sigma^A \otimes \sigma^B \,|\, 00 \rangle = \langle 1 \,|\, \sigma^A \,|\, 0 \rangle \langle 1 \,|\, \sigma^B \,|\, 0 \rangle$) makes an operator (e.g. σ^A or σ^B) work on the wavefunction on its space only. This is exactly the case of measurement (or vector component decomposition) on different axes. Third, the orthogonal condition of basis wavefunctions mimics the measurement of the projection onto the orthogonal axes, e.g. $\langle 0 \,|\, \sigma^A \,|\, 0 \rangle = +1$, $\langle 1 \,|\, \sigma^B \,|\, 1 \rangle = -1$, and $\langle 1 \,|\, \sigma^A \,|\, 0 \rangle = 0$. Fourth, the space (or axis) separation is consistent with the concept of correlation. For example, since Alice measures on the x-axis, only the x-component of the measurement by Bob is relevant to the correlation calculations. This is manifested by $\langle 0 \,|\, \sigma^A \,|\, 0 \rangle \langle 0 \,|\, \sigma^B \,|\, 0 \rangle = \langle 0 \,|\, \sigma^A_x \,|\, 0 \rangle \langle 0 \,|\, \sigma^B_x \,|\, 0 \rangle$. Finally, the normalized wavefunction automatically normalizes the calculated expected value so that it fits the requirement of correlation.

If we use other entangled wavefunctions to perform similar calculations, we would arrive at essentially the same results but with a negative sign for some wavefunctions. For example, with $\phi = (\,|01\rangle + |10\rangle)/\sqrt{2}$, we find:

$$\langle \sigma^A \sigma^B \rangle = \langle \phi \,|\, \sigma^A \otimes \sigma^B \,|\, \phi \rangle = -\cos\beta\,.$$

The above result is equivalent to (16) with $\theta_0 = \pi$. This is not surprising as this wavefunction indicates a negative correlation. If we use a wavefunction of separable states to calculate the expected joint measurement, we would find a value of zero. This is expected because there is no correlation between separable states.

If the measurement axes change randomly, we cannot put a vector on either the A or B axis. In this case, the QM derivation of the joint detection rate involves a projection process similar to that used in Fig. 2. Using a matrix presentation, we can express the projection of a vector pointing to (x_1, y_1) onto a specified axis of angle θ as follows:

$$\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \\ = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}. \tag{22}$$

In the above equation, if we let θ be the angle of the B axis with respect to the x-axis, i.e. $\theta - \theta_b$ in Fig. 2, and let $x_1 = \overrightarrow{a_x}/\sqrt{2}$ and $y_1 = \overrightarrow{a_y}/\sqrt{2}$, we can obtain the same result as in (19) and (20).

The matrix in (22) is called a projection matrix [28], as it projects a vector onto the axis of angle θ and gives the components of the projection:

$$Q(\theta) = \begin{pmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}.$$

Using the above projection matrix and an entangled wavefunction (e.g. $\phi = (\,|01\rangle + |10\rangle)/\sqrt{2}$), we can calculate the probability of joint measurement as:

$$\begin{aligned} p_{AB} &= \langle \psi \,|\, Q^A \otimes Q^B \,|\, \psi \rangle = 0.5\cos^2(\theta_A - \theta_B) \\ &= 0.25\,[\cos 2(\theta_A - \theta_B) + 1]\,. \end{aligned}$$

Since $(\theta_A - \theta_B)$ is the angle of the orientations of detectors, the above result is exactly the same as (18) or (21) that we derived in the statistical approach. The identical result is apparently because the same projection process works in both approaches.

5 Statistical explanation of two-particle Bell tests

Many Bell test experiments are based on the coincidence rate of particle pairs, but a handful of researchers (e.g. [9, 16, 20, 32, 39, 40] have conducted experiments on correlations of 3 or more particles. Multi-particle correlation can be achieved by special designs of experimental setup to obtain specific quantum states (e.g. [9, 32]) or by exploiting the coherent states of Bose-Einstein condensate (e.g. [16, 20]). The statistical foundation of multi-particle correlation is the same as that for particle pairs, so this paper focus on two-particle correlation. Even though we confine our scope to two-particle Bell tests, there still are copious experiments. This section selects only some representative experiments and puts them into two groups: the polarization experiments of entangled photon pairs and non-polarization experiments based on light phase correlation.

5.1 Polarization experiments

Among numerous Bell test using polarization of photon pairs, we consider only two influential papers by Aspect *et al* [5, 6]. Like most experiments on the Bell tests, Aspect *et al* [5, 6] utilized the derivation of Clauser *et al* [4] for an experimentally applicable quantum mechanical prediction for the counting rates of coincidence. The starting point of their derivation is a probability formula:

$$\begin{aligned} P(a,b) = {} & w[A(a)_+, B(b)_+] - w[A(a)_+, B(b)_-] \\ & - w[A(a)_-, B(b)_+] + w[A(a)_-, B(b)_-] \end{aligned}$$

where w means the probability weighting of each outcome in total emission counts R_0, with:

$$\begin{aligned} R_0 = {} & [A(a)_+, B(b)_+] + [A(a)_+, B(b)_-] \\ & + [A(a)_-, B(b)_+] + [A(a)_-, B(b)_-] \end{aligned}$$

$$w[A(a)_+, B(b)_+] = [A(a)_+, B(b)_+]/R_0,\ etc.$$

The above equation is a manifest that the net correlation (positive correlation $[A(a)_+, B(b)_+] + [A(a)_-, B(b)_-]$ minus negative correlation $[A(a)_+, B(b)_-] + [A(a)_-, B(b)_+]$) in terms of total counts R_0. This equation is consistent with our derivation of joint detection rate presented in Section 3: the net correlation in (17) is calculated by integrating E_{AB} over the angle $0-2\pi$ while the total counts is obtained by integrating the unit spin vector over the same range. Due to the same foundation for derivation, the resulting (18) is unsurprisingly the same as that obtained by Clauser *et al* [4] and used by Aspect *et al* [5, 6]. Since the joint detection rate derived from both statistical and quantum approaches is identical, the explanation on the results of Aspect *et al* [5, 6] will be very similar, so we omit this explanation but examine the maximum violation angle derived from quantum mechanics and confirmed by experiments.

Using the coplanar vectors (shown in Fig. 4) introduced by Clauser and Shimony [28] and Aspect *et al* [5] to present the settings of the Bell test experiments, we can derive the same results as the quantum prediction of the Bell test, but without invoking any wavefunctions.

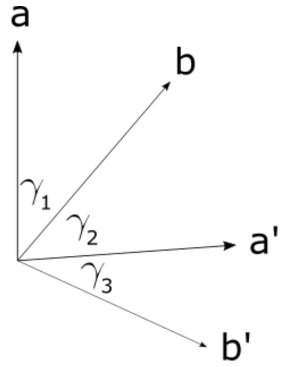


Fig. 4: Coplanar vectors presentation of Bell test settings

In Fig. 4, vectors a, a', b, and b' represent the direction of the spin detectors, and the angles between them are displayed on the graph. For simplicity of presentation, we assume all vectors are of unit modulus and angles γ_1, γ_2, and γ_3 are positive and less than π (for any angle θ greater than π, we can rewrite it as $2\pi - \theta$). Applying the spin correlation results in (16) derived in Section 3 to a case of positively entangled particles (i.e. $\theta_0 = 0$), we can obtain the experimental results as follows:

$$\begin{aligned} E(a,b) &= \cos\gamma_1 \\ E(a,b') &= \cos(\gamma_1+\gamma_2+\gamma_3) \\ E(a',b) &= \cos\gamma_2 \\ E(a',b') &= \cos\gamma_3\,. \end{aligned}$$

The theoretical results for the Bell tests should be:

$$\begin{aligned} E_{BT} &= E(a,b) - E(a,b') + E(a',b) + E(a',b') \\ &= \cos\gamma_1 - \cos(\gamma_1+\gamma_2+\gamma_3) + \cos\gamma_2 + \cos\gamma_3\,. \end{aligned}$$

Applying the first and second order conditions of maximization (minimization) for the above equation, we know that E_{BT} reaches the maximum or minimum when:

$$\sin\gamma_1 = \sin\gamma_2 = \sin\gamma_3 = \sin(\gamma_1+\gamma_2+\gamma_3)\,.$$

If γ_1, γ_2 and γ_3 are less than $\pi/2$, the condition of maximum/minimum value necessitates that $\gamma_1 = \gamma_2 = \gamma_3 = \gamma$ and $\sin\gamma = \sin 3\gamma$. With some trigonometric manipulations, from $\sin\gamma = \sin 3\gamma$ we can have $\sin\gamma\,(4\cos 2\gamma\text{-}1) = \sin\gamma$, or $\gamma = \pi/4$.

Similarly, if γ_1, γ_2 and γ_3 are greater than $\pi/2$ (they are less than π as we assumed before for simplicity), we can obtain $\gamma = 3\pi/4$.

If some angles are less than $\pi/2$, but some are greater than $\pi/2$, we obtain no satisfying solution. For example, if γ_1 and γ_2 are less than $\pi/2$, but γ_3 is greater than $\pi/2$, from $\sin\gamma_1 = \sin\gamma_2 = \sin\gamma_3$, we can infer that $\gamma_1 = \gamma_2$ and $\gamma_3 = \pi - \gamma_2$, so $\sin\gamma_1 = \sin(\gamma_1+\gamma_2+\gamma_3) = \sin(\gamma_1+\pi) = -\sin\gamma_1$, or $\gamma_1 = \gamma_2 = \gamma_3 = 0$. This contradicts our assumption of positive angles and presents a trivial case where all 4 settings coincide.

To sum up, from the first and second order condition we reveal that the maximum and minimum value of E_{BT} occurs at $\gamma = \pi/4$ and $\gamma = 3\pi/4$, respectively. If $\gamma = \pi/4$, we have:

$$E_{max} = \cos\pi/4 - \cos 3\pi/4 + \cos\pi/4 + \cos\pi/4 = 2\sqrt{2}\,.$$

If $\gamma = 3\pi/4$, we have:

$$E_{min} = \cos 3\pi/4 - \cos 9\pi/4 + \cos 3\pi/4 + \cos 3\pi/4 = -2\sqrt{2}\,.$$

As a result, we obtain the same results as the quantum prediction:

$$\left|E(a,b) - E(a,b') + E(a',b) + E(a',b')\right| \le 2\sqrt{2}\,.$$

It is worth mentioning that the above derivation shows that the maximum violation of the Bell inequality occurs at the setting $\gamma = \pi/4$ or $\gamma = 3\pi/4$, $E = \pm 2\sqrt{2}$. This seems in conflict with the results of Aspect *et al* [5, 6], where the maximum violation of the Bell inequality occurred at $\theta = \pi/8$, or $\theta = 3\pi/8$.

In fact, this difference highlights the different cases of spin and polarization. Our derivation is based on spin detection. As we discussed in Section 3, the angle must be adjusted when applying (16) and (17) to polarization experiments. In most Bell test experiments using light, including Aspect *et al* [5, 6], a count of photon detection is recorded as +1 and no detection is recorded as -1. As such, if the angle between the polarizer and the polarization of light is $\theta = \pi/2$, the most likely outcome is no detection or -1. We can express the result as $\cos 2\theta = \cos\pi = -1$. It is apparent that one needs to double the angle in the experiment to obtain a result that is consistent with experimental record. On the other hand, our derivation based on spin assumes that a count of photon detection is recorded as +1 and no detection is recorded as 0. If the angle between the polarizer and the polarization of light is $\gamma = \pi/2$, the most likely outcome is no detection or 0. We can express the result as $\cos\gamma = \cos\pi/2 = 0$. This recorded value is equivalent to the case of $\theta = \pi/4$ in Aspect *et al* [5, 6]. From this we can infer that the angle γ used for spin examples in the present paper is equivalent to twice the angle θ used in Aspect *et al* [5, 6], i.e. $\gamma = 2\theta$. As a result, the angles for maximum violation of the Bell inequality in Aspect *et al* [5, 6] will be half the value as in our derivation.

5.2 Interferometry Bell tests

There are Bell tests that examine the correlations between variables other than polarization. One type of research focuses on the phase correlation (e.g. [19, 34–36]). This type of experiment creates a pair of photons of the same phase and lets them pass through phase shifters and a distance of different lengths, then detects the phase difference at a Michelson interferometer. The experiments are based on the theoretical prediction of Franson [37] which, based on the phase difference of wavefunctions caused by time difference, developed a similar prediction as (18) in the present paper. Using a classical wave theory of light and joint intensity, one can also obtain an equivalent result.

For simplicity, we combine the electrical and magnetic components of a light field, so the normalized light field of a photon pair of the same initial phase at position x and time t can be expressed as:

$$E = \cos(\theta + kx - \omega t)$$

where θ is the initial phase of the photon pair at the source, k is wave vector, ω is angular frequency.

Assume that photon A will be added a phase θ_a by a phase shifter (we use only one phase shifter for simplicity) and, meanwhile, photon B will be added a phase $\theta_b = \omega\Delta t$ due to the different time or distance travelled. The light fields of the pair become:

$$E_A = \cos(\theta + kx - \omega t + \theta_a)$$

$$E_B = \cos(\theta + kx - \omega t + \theta_b)\,.$$

Although this type of experiments use the joint intensity as measurement, as we discussed in Section 3, we cannot calculate the correlation of light intensity by directly multiplying the intensities of light field because the changes in intensity are not independent. Since the light phases and thus the light fields are correlated, the joint intensity needs to be calculated from light field correlation:

$$\begin{aligned} E_{AB} &= E_A E_B \\ &= \cos(\theta + kx - \omega t + \theta_a)\cos(\theta + kx - \omega t + \theta_b) \\ &= 0.5\,[\cos(2\theta + 2kx - 2\omega t + \theta_a + \theta_b) + \cos(\theta_a - \theta_b)]\,. \end{aligned}$$

The initial phase of photon pair θ can change randomly, so the item related to θ in the above equation will net out to zero (by integrating E_{AB} over θ in the range of $0 - 2\pi$). As a result, the above equation becomes:

$$E_{AB} = 0.5\cos(\theta_a - \theta_b)\,.$$

As such, the joint intensity can be calculated as:

$$I_{AB} = E^2_{AB} = 0.25\cos^2(\theta_a - \theta_b) = 0.125\,[\cos 2(\theta_a - \theta_b) + 1]\,.$$

This result is equivalent to the quantum prediction in Franson [37, Eq. (16)] or Brendel *et al* [34, Eq. (4)]. From the above derivation we can conclude that the light intensity difference stems from the phase difference caused by phase shifter and by different travel time. Probability law also works in this case because it ensures that the initial random phase of photon pairs have no impact on the interferometry results.

By examining representative experiments, we can conclude that the violation of the Bell inequality is caused by the correlation in source particles as well as the physical relationship between the spin/polarization angle and its component on detection axes, or between the phase of electromagnetic wave and the light field correlation. With varying detection conditions (i.e. random changes in detection angles or adding arbitrary phases), probability law can still maintain the correlation of source particles. This leads to the violation of the Bell inequality and the correct statistical predictions, which are consistent with experimental outcomes.

6 Conclusions

The paper presents statistical predictions of two-particle Bell tests, which are equivalent to, but more general than, the QM predictions. By comparing the statistical and QM approaches, the paper shows that probability law is at the heart of both approaches. The statistical presentation of two-particle Bell tests in this paper has far-reaching implications.

First, it can improve our understanding of quantum mechanics and help to demystify it. Although the concepts of superposition and entanglement are widely accepted among physicists, the explanation of these concepts is difficult and thus causes significant misunderstanding. The statistical interpretation of the Bell tests shows that the superposition of entangled states in the two-particle Bell test is nothing more than statistical correlation between states. For the correlated particles at the polarizer or spin detector, probability law can maintain the correlation through the expected value, so there is no need for communication (let alone faster-than-light or instantaneous communication) between different locations in the Bell experiments. As quantum entanglement is explained by probability law, the Bell test results and quantum mechanics are no longer mysterious.

Second, it has significant implications for quantum computing, which relies on quantum entanglement. Since the quantum entanglement phenomenon results from probability law, statistical noise is a natural and unavoidable part of quantum computing. Understanding the nature of this noise may shed light on how to improve the signal-to-noise ratio and thus is crucial to the success of quantum computing.

Third, the paper pinpoints the cause for the violation of the Bell inequality and thus explains why the local hidden variable theory is wrong. Although numerous Bell tests reject the local hidden variable theory and support quantum mechanics, they have not shed any light on why the former is wrong and the latter is right. This paper shows that the key lies in probability law, which underpins the Bell test results. Because probability law is universal, if we regard the statistical mechanism (which causes statistical variation around the mean) as a "hidden variable", it is not a local one but a global one. The local hidden variable theory misrepresents this nature and thus fails. It is also this global law that leads to the correct prediction from quantum mechanics.

Last but not least, the paper may stimulate a reassessment of the role of determinism and realism. Broadly, the experimental results on the Bell inequality are interpreted as being a rejection of determinism or local realism, and an embracing of randomness. While this paper highlights the importance of randomness and probability law, it does not totally reject determinism and realism. In the Bell tests, probability law works only when the particles arrive at and interact with the detector (polarizer or spin analyzer) – it plays no role before and after. When probability law is not in action, it is determinism, realism and logic that describe the behaviour of the particles. In other words, both randomness and realism play important roles in our understanding of physics.

Acknowledgements

The author thanks Sean Hodgman and David Shin for useful discussions and comments.

Received on May 20, 2021

References

1. Einstein E., Podolsky B., Rosen N. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 1935, v. 47, 777–780.
2. Bohm D. A Suggested Interpretation of the quantum theory in terms of "hidden" variables. *Physical Review*, 1952, v. 85, 169–193.
3. Bell J. On the Einstein Podolsky Rosen paradox. *Physics*, 1964, v. 1 (3), 195–200.
4. Clauser J. F., Horne M. A., Shimony A., Holt R. A. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 1969, v. 23, 880.
5. Aspect A., Grangier P., Roger G. Experimental tests of realistic local theories via Bell's theorem. *Physical Review Letters*, 1981, v. 47 (7), 460–463.
6. Aspect A., Grangier P., Roger G. Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A new violation of Bell's Inequalities. *Physical Review Letters*, 1982, v. 49 (2), 91–94.
7. Tittel W., Brendel J., Gisin B., Herzog T., Zbinden H., Gisin N. Experimental demonstration of quantum-correlations over more than 10 kilometers. *Physical Review A*, 1998, v. 57 (5), 3229–3232.
8. Weihs G., Jennewein T., Simon C., Weinfurter H., Zeilinger A. Violation of Bell's inequality under strict Einstein locality conditions. *Physical Review Letters*, 1998, v. 81 (23), 5039–5043.
9. Pan J. W., Bouwmeester D., Daniell M., Weinfurter H., Zeilinger A. Experimental test of quantum nonlocality in three-photon GHZ entanglement. *Nature*, 2000, v. 403 (6769), 515–519.
10. Rowe M. A., Kielpinski D., Meyer V., Sackett C. A., Itano W. M., Monroe C., Wineland D. J. Experimental violation of a Bell's inequality with efficient detection. *Nature*, 2001, v. 409 (6822), 791–794.
11. Gröblacher S., Paterek T., Kaltenbaek R., Brukner S., Zukowski M., Aspelmeyer M., Zeilinger A. An experimental test of non-local realism. *Nature*, 2007, v. 446 (7138), 871–875.
12. Salart D., Baas A., van Houwelingen J., Gisin N., Zbinden H. Spacelike Separation in a Bell Test Assuming Gravitationally Induced Collapses. *Physical Review Letters*, 2008, v. 100 (22), 220404.
13. Ansmann M., Wang H., Bialczak R., Hofheinz M., Lucero E., Neeley M., O'Connell A. D., Sank D., Weides M., Wenner J., Cleland A. N., Martinis J. M. Violation of Bell's inequality in Josephson phase qubits. *Nature*, 2009, v. 461, 504–506.
14. Giustina M., Mech A., Ramelow S., Wittmann B., Kofler J., Beyer J., Lita A., Calkins B., Gerrits T., Nam S. W., Ursin R., Zeilinger A. Bell violation using entangled photons without the fair-sampling assumption. *Nature*, 2013, v. 497 (7448), 227–230.
15. Christensen B. G., McCusker K. T., Altepeter J., Calkins B., Gerrits T., Lita A., Miller A., Shalm L. K., Zhang Y., Nam S. W., Brunner N., Lim C. C. W., Gisin N., Kwiat P. G. Detection-Loophole-Free Test of Quantum Nonlocality, and Applications. *Physical Review Letters*, 2013, v. 111 (7448), 130406.
16. Schmied R., Bancal J.-D., Allard B., Fadel M., Scarani V., Treutlein P., Sangouard N. Bell correlations in a Bose-Einstein condensate. *Science*, 2016, v. 352 (6284), 441–444.
17. Handsteiner J., Friedman A. S., Rauch D., Gallicchio J., Liu B., Hosp H., Kofler J., Bricher D., Fink M., Leung C., Mark A., Nguyen H. T., Sanders I., Steinlechner F., Ursin R., Wengerowsky S., Guth A. H., Kaiser D. I., Scheidl T., Zeilinger A. Cosmic Bell Test: Measurement Settings from Milky Way Stars. *Physical Review Letters*, 2017, v. 118 (6), 060401.
18. BIG Bell Test Collaboration. Challenging local realism with human choices. *Nature*, 2018, v. 557 (7704), 212–216.
19. Hensen B., Bernien H., Dréau A. E., Reiserer A., Kalb N., Blok M. S., Ruitenberg J., Twitchen R. F. L. Vermeulen & Schouten R. N. & Abellán C. & Amaya W. & Pruneri & Mitchell M. W. & Markham D. J., Elkouss D., Wehner S., Taminiau T. H., Hanson R. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 2015, v. 526, 682–686.
20. Shin D. K., Henson B. M., Hodgman S. S., Wasak T., Chwedenczuk J., Truscott A. G. Bell correlations between spatially separated pairs of atoms. *Nature Communications*, 2019, v. 104447.
21. Thenabadu M., Cheng G.-L., Pham T. L. H., Drummond L. V., Rosales-Zárate L., Reid M. D. Testing macroscopic local realism using local nonlinear dynamics and time settings. *Physical Review A*, 2020, v. 102, 022202.
22. Bell J. Introduction to the hidden variable question. In: d'Espagnat B., ed. Foundations of Quantum Mechanics. 1971, pp. 171–181.
23. Myrvold W., Genovese M., Shimony A. Bell's Theorem. In: Edward N. Zalta, ed. The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, 2019.
24. Pierce D. T., Celotta R. J., Kelley M. H., Unguris J. Electron spin polarization analyzers for use with synchrotron radiation. *Nuclear Instruments and Methods in Physics Research A*, 1988, v. 266, 550–559.
25. Getzlaff M., Heidemann B., Basmann J., Westphal C., Schonhense G. A variable-angle electron spin polarization detection system. *Review of scientific instruments*, 1998, v. 69 (11), 3913–3923.
26. Yu D., Math C., Meier M., Escher M., Rangelov G., Donath M. Characterisation and application of a SPLEED-based spin polarisation analyser. *Surface Science*, 2007, v. 601, 5803–5808.
27. Winkelmann A., Hartung D., Engelhard H., Chiang C.-T., Kirschner J. High efficiency electron spin polarization analyser based on exchange scattering at Fe/W(001). *Review of Science Instruments*, 2008, v. 79, 083303.

28. Clauser J. F., Shimony A. Bell's theorem: experimental tests and implications. *Reports on Progress in Physics*, 1978, v. 41, 1881.
29. Bell J. Bertlmann's socks and the nature of reality. *Journal De Physique*, 1981, v. C2, 41–62.
30. Fine A. Hidden variables, joint probability, and the Bell inequalities. *Physical Review Letters*, 1982, v. 48 (5), 291–295.
31. Ou Z. Y., Mandel L. Violation of Bell's inequality and classical probability in a two-photon correlation experiment. *Physical Review Letters*, 1988, v. 61 (1), 50–53.
32. Bouwmeester D., Pan J. W., Daniell M., Weinfurter H., Zeilinger A. Observation of three-photon Greenberger-Horne-Zeilinger entanglement. *Physical Review Letters*, 1999, v. 82 (7), 1345–1349.
33. Greenberger D. M., Horne M. A., Zeilinger A. Going beyond Bell's theorem. In: Kafatos M., ed. Bell's Theorem, Quantum Theory and Conceptions of the Universe. Kluwer Academics Dordrecht, The Netherlands, 1989, pp. 69–72.
34. Brendel J., Mohler E., Martienssen W. Time-resolved dual-beam two-photon interferences with high visibility. *Physical Review Letters*, 1991, v. 66 (9), 1142–1145.
35. Rarity J. G., Tapster P. R. Experimental violation of Bell's inequality based on phase and momentum. *Physical Review Letters*, 1990, v. 64 (21), 2495–2498.
36. Tittel W., Brendel J., Zbinden H., Gisin N. Violation of Bell inequalities by photons more than 10km apart. *Physical Review Letters*, 1998, v. 81 (17), 3563–3566.
37. Franson J. D. Bell inequality for position and time. *Physical Review Letters*, 1989, v. 62 (19), 2205–2208.
38. Freedman S. J., Clauser J. F. Experimental test of local hidden-variable theories. *Physical Review Letters*, 1972, v. 28, 938–941.
39. Lanyon B. P., Zwerger M., Jurcevic P., Hempel C., Dür W., Briegel H. J., Blatt R., Roos C. F. Experimental violation of multipartite Bell inequalities with trapped ions. *Physical Review Letters*, 2014, v. 112 (10), 100403.
40. Eibl M., Gaertner S., Bourennane M., Kurtsiefer C., Żukowski M., Weinfurter H. Experimental observation of four-photon entanglement from parametric down-conversion. *Physical Review Letters*, 2003, v. 90, 200403.

A Quantitative Description of Atmospheric Absorption and Radiation at Equilibrium Surface Temperature

Y. C. Zhong

ERICHEN Consulting, Queensland, Australia. E-mail: drzhong88@yahoo.com

An analytical theory is proposed for the earth-atmosphere system at its equilibrium surface temperature, 289.16 K. A non-linear relation is formulated between atmospheric absorption and atmospheric radiation by modifying Kirchhoff's law on thermal radiation. For the first time, the Global Energy Balance can be realized in a wide range of atmospheric absorptivity, transmittance, and surface emissivity. It is revealed that atmospheric radiation becomes negative once the atmospheric absorptivity is below its threshold value. It is proven that the upward cumulative long-wave atmospheric radiation spontaneously increases from 3.8 W m^{-2} to 199.4 W m^{-2} as the long-wave atmospheric absorptivity increases from 0.4 to 1.0 whilst the long-wave atmospheric transmittance decreases from 0.6 to 0.1.

1 Introduction

For over a century, many attempts have been made to balance the global energy budget, both at the top of the atmosphere (TOA) and at the Earth's surface [1]. It is known that the lack of precise knowledge of the surface energy fluxes profoundly affects the ability to study climate change [2]. In fact, the power equation at the surface remains unbalanced as the uncertainty in the net energy flux between the surface and the atmosphere is over 17 W m^{-2} [3]. To date, many static explanations for the global energy balance have been confined to using one set of fixed parameters to describe atmospheric absorption and radiation [2], whereas the taken-for-granted Kirchhoff's law at the core of the radiative transfer description of atmospheric absorption and radiation seems theoretically invalid [4].

In this paper, several thermodynamic variables of theoretical importance are redefined to formulate the basic equations, including those previously treated as constants. By continuously mapping the surface emissivity and longwave (LW) atmospheric absorptivity, several coupled quadratic equations are derived and simultaneously solved, which are in quantitative agreement with the latest experimental observations. In light of these new findings, implications for some fundamental issues in climate studies are briefly discussed.

2 Theory

In general, the thermodynamic variables in the atmosphere-surface system are dependent and should be described in coupled equations.

2.1 Outgoing longwave radiation and surface radiance

It is known that the total power balance at the TOA can be written as

$$\pi R^2 S\,(1-r) = 4\pi R^2 I_{LW}^{\uparrow} \tag{1}$$

where S is the solar constant, R the radius of the Earth, r the effective reflectivity of the Earth at the TOA, including the SW solar radiation reflected at the surface and then transmitted upward to the TOA, and $I_{LW}^{\uparrow}$ denotes the outgoing LW radiation (OLR) into outer space. From (1),

$$I_{LW}^{\uparrow} = \frac{S\,(1-r)}{4}\,. \tag{2}$$

Notice that OLR is merely determined by the albedo and the solar constant.

By treating the Earth as a graybody, the surface radiation can be obtained from the Stefan-Boltzmann law,

$$I_E = \varepsilon_E \sigma T^4\,, \tag{3}$$

where ε_E is defined as the Earth's mean surface emissivity, and T is the equilibrium mean surface temperature. In general, ε_E is to be treated as a thermodynamic variable in this study, although it has been often approximated as unity so far.

2.2 Modification of Kirchhoff's Law

In theory, the upward *cumulative* atmospheric absorption at any altitude can be calculated using the line-by-line method provided all of the relevant lineshape functions are known. At the TOA, the total LW atmospheric absorption can be expressed as

$$A_{LW} = \iint_0^\infty \alpha_\lambda\,(T_A)\,\rho\,(z)\,I_E\,(\lambda, z)\,d\lambda\,dz\,, \tag{4}$$

where α_λ is the spectral absorptivity of the atmosphere, predominately determined by water vapor, T_A is the atmospheric temperature at at a given altitude, ρ is the air density, $I_E\,(\lambda, z)$ represents the attenuated surface LW emission spectra at different altitudes. Naturally, α_λ represents both the resonant and continuum absorption by air molecules detected under continuous excitation [5,7]. Note that α_λ is scaled by the Planck function $B\,(\lambda, T)$ with its maximum at the center of the atmospheric window near 10 μm.

To proceed further, an effective LW cumulative atmospheric absorptivity, a_{LW}, at the TOA can be introduced

$$A_{LW} = a_{LW} I_E \,. \tag{5}$$

Obviously, the maximum LW atmospheric absorption is I_E when $a_{LW} = 1$.

Similarly, SW atmospheric absorption can be written as

$$A_{SW} = a_{SW} \left[1 + r_{SE} \left(1 - a_{SW}\right)\right] I^{\downarrow}_{SW} (TOA) \,. \tag{6}$$

where $I^{\downarrow}_{SW}(TOA)$ in the actual downward SW solar radiation at the TOA by subtracting the reflected SW solar radiation at the TOA, r_{SE} is defined the SW surface reflectivity. In this study, the SW atmospheric absorption is fixed.

Using Kirchhoff's law, it would appear that $\alpha = \varepsilon$, where α and ε are the spectral absorptivity and the emissivity of a non-scattering medium, respectively. Nevertheless, it seems unrealistic to expect that atmospheric radiation is equal to atmospheric absorption. So far, many attempts have been made at *ab initio* calculation of atmospheric radiation based on Schwarzschild's equation with the Planck function and an effective emissivity, but the results seem over-simplified. Besides, it has been revealed that Kirchhoff's law is problematic and should not be considered as a basic law [4].

In this paper, it is postulated that the fraction, denoted by β, of upward cumulative atmospheric radiation (UCAR), is proportional to the LW atmospheric absorptivity

$$a_{LW} = \gamma\beta \tag{7}$$

where γ denotes the proportionality factor that is used to parameterize the rest of the unclear dependence during radiative transfer in the atmosphere. In effect, (7) can be considered as a modified Kirchhoff's law for atmospheric radiation. In the absence of internal reflection, it would appear the sum of the LW atmospheric absorptivity and the LW atmospheric transmittance, τ_{LW}, is unity.

$$\tau_{LW} = 1 - a_{LW} \,. \tag{8}$$

Substituting (7) into (8) yields

$$\tau_{LW} = 1 - \gamma\beta \,. \tag{9}$$

It is shown in this study, however, that (8) and (9) are invalid in the presence of atmospheric radiation which is empowered by atmospheric absorption and other non-radiative energy fluxes.

2.3 Formulation for power balance conditions

To derive the power balance equation at the surface, that ensures the net energy flux at surface is exactly zero at thermal equilibrium, the net downward energy flux (NDEF) is denoted as N_0. Thus the power balance equation at the surface can be simply written as

$$N_0 = I_E \,. \tag{10}$$

As the downward SW solar radiation into the surface $I^{\downarrow}_{SW}(0)$ is known, it can be taken away from N_0 and explicitly expressed in the power balance condition,

$$N + I^{\downarrow}_{SW}(0) = I_E \,, \tag{11}$$

where N represents the NDEF when $I^{\downarrow}_{SW}(0)$ is excluded from N_0, *viz.*

$$N = N_0 - I^{\downarrow}_{SW}(0) \,. \tag{12}$$

Note that (11) and (12) are equivalent irrespective of the value of $I^{\downarrow}_{SW}(0)$.

At the TOA, the power balance equation for OLR can be expressed as,

$$I^{\uparrow}_{LW} = \tau_{LW} I_E + I^{\uparrow}_{A} \tag{13}$$

where τ_{LW} is the LW atmospheric transmittance, $I^{\uparrow}_{A}$ is the UCAR that can escape from the atmosphere into space. It is to be shown that the upward LW radiation at the TOA is a constant.

2.4 Formulation for atmospheric radiation

In the absence of the physical surface underneath the atmosphere while the LW radiation were still available, the upward LW atmospheric radiation at the TOA can be obtained by assuming it is proportional to the total atmospheric absorption without invoking Stefan-Boltzmann law.

$$I^{\uparrow}_{A} = \beta \left(A_{LW} + A_{SW}\right) \,. \tag{14}$$

The two absorption terms in (14) belong to, respectively, the one-way cumulative LW atmospheric absorption from the surface radiation A_{LW}, and the two-way cumulative SW atmospheric absorption from the solar radiation A_{SW}. In this hypothetical case, those non-radiative energy exchange processes are absent.

Similarly, the downward cumulative atmospheric radiation (DCAR) at the bottom of the atmosphere, can be derived

$$I^{\downarrow}_{A} = \left(1 - \beta\right) \left(A_{LW} + A_{SW}\right) \,. \tag{15}$$

Adding (14) and (15) yields,

$$I^{\uparrow}_{A} + I^{\downarrow}_{A} = a_{LW} I_E + A_{SW} \,, \tag{16}$$

which is simply an energy conservation statement.

In reality, however, the bottom of the atmosphere is physically in contact with the Earth's surface, hence the thermal energy exchange, in addition to radiation, is inevitable. As a result, (14)–(16) should be modified accordingly. Specifically, a portion of the total energy absorbed by the atmosphere must be used to achieve and maintain the thermal equilibrium in the atmosphere-surface system, as required by (11), which is exactly equal to N. Thus we have,

$$I^{\uparrow}_{A} = \beta \left(A_{LW} + A_{SW} - N\right) \,, \tag{17}$$

$$I_A^{\downarrow} = (1-\beta)\,(A_{LW} + A_{SW} - N)\;, \tag{18}$$

$$I_A^{\uparrow} + I_A^{\downarrow} = a_{LW} I_E + A_{SW} - N\,. \tag{19}$$

Note that (19) predicts that the total atmospheric radiation can either be zero or negative if the total atmospheric absorption is equal to or less than N, respectively. Here, (19) is to be used as the criterion to quantitatively determine the eventuating total atmospheric radiation, $I_A^{\uparrow} + I_A^{\downarrow}$, which, in turn, allows calculation for other climate variables, such as LW atmospheric transmittance and the net downward energy flux (NDEF).

2.5 Corollary

Substituting (5), (7), and (17) into (13), the power balance condition at the TOA can be rewritten as a quadratic function of the UCAR fraction β,

$$\tau_{LW} = -\gamma\beta^2 - \left(\frac{A_{SW} - I_E + I_{SW}^{\downarrow}(0)}{I_E}\right)\beta + \frac{I_{LW}^{\uparrow}}{I_E}\,, \tag{20}$$

with its y-intercept close to 0.6, which is determined by the ratio of OLR to the surface radiation. Note that (20) indicates that LW atmospheric transmittance is not unity in the absence of UCAR, as derived from (9) and shown in Fig. 1, due to the contribution of SW absorption by the atmosphere.

Substituting (7) into (20), we obtain the dependence of LW atmospheric transmittance on the LW atmospheric absorptivity,

$$\tau_{LW} = -\frac{a_{LW}^2}{\gamma} - \left(\frac{A_{SW} - I_E + I_{SW}^{\downarrow}(0)}{\gamma I_E}\right) a_{LW} + \frac{I_{LW}^{\uparrow}}{I_E}\,, \tag{21}$$

which indicates that the relation between LW transmittance and LW absorptivity is not linear, but quadratic, as shown in Fig. 2. As a result, the well-known linear relation between τ_{LW} and a_{LW}, (9), should be replaced by (21). To obtain the analytical formula for the atmospheric radiation that satisfies energy conservation law, substituting (5) and (7) into (17) yields a quadratic equation for UCAR,

$$I_A^{\uparrow} = \gamma I_E \beta^2 + \left(A_{SW} - I_E + I_{SW}^{\downarrow}(0)\right)\beta\,. \tag{22}$$

Dividing (17) by (18) and then substituting the result into (22) yields,

$$I_A^{\downarrow} = (1-\beta)\left(\gamma I_E \beta + A_{SW} - I_E + I_{SW}^{\downarrow}(0)\right)\,. \tag{23}$$

3 Calculated results

Based on the latest experimental data used in [7] and [8], as shown in Table 1, all of the numerical calculations are based on solving the coupled quadratic equations, (20) to

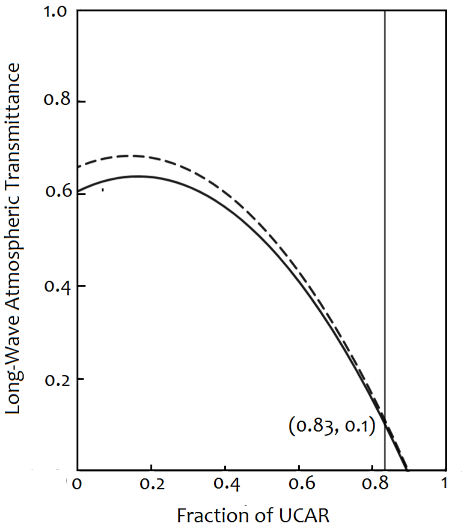


Fig. 1: Dependence of LW atmospheric transmittance τ_{LW} on the fraction of UCAR at the TOA, calculated from (20) assuming the surface emissivity is 1.0 (solid curve) and 0.92 (dashed curve). The coordinate (0.83, 0.1) represents the maximum β at $\tau_{LW} = 0.1$, used in this study.

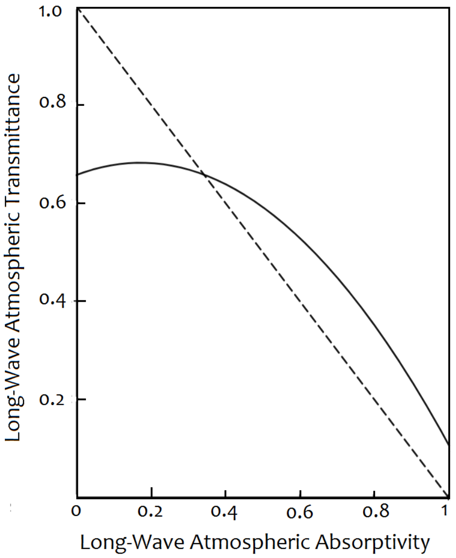


Fig. 2: Dependence of LW atmospheric transmittance on LW atmospheric absorptivity at the TOA, obtained from (21) in this study (solid curve) and from (8) (dashed line).

surface mean temperature	289.16	K
albedo	0.2985	
solar constant	1365.2	W m^{-2}
reflected solar radiation at TOA	101.9	W m^{-2}
SW atmospheric absorption	78	W m^{-2}
surface solar SW radiation	161	W m^{-2}

Table 1: The observed data used in [7] and this study.

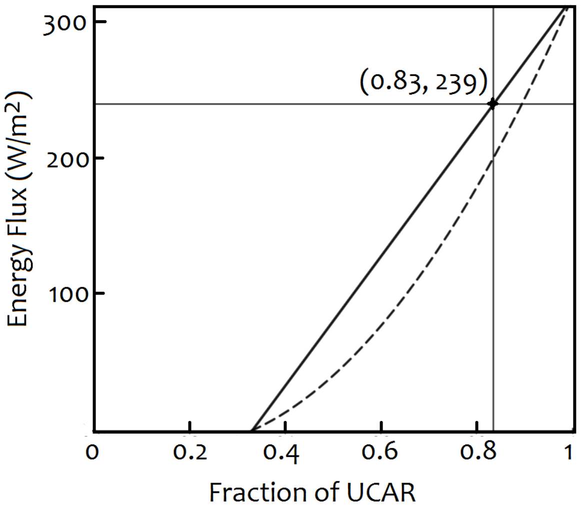


Fig. 3: Dependence of the total LW atmospheric radiation (solid line) and LW UCAR (dashed curve) on the fraction of UCAR the at the TOA, calculated from (19) and (22), respectively.

(24). A wide range of different values for surface emissivity and LW atmospheric absorptivity are considered. Specifically, the proportionality γ-factor in (7) is first determined by using the LW atmospheric transmittance $\tau_{LW} = 0.1009$ at the surface emissivity $\varepsilon_E = 1$ and then by maximizing the LW atmospheric absorptivity to $a_{LW} = 100\%$. This operation is equivalent to first assuming the atmospheric transmittance becomes its minimum whilst the LW atmospheric absorption reaches to its maximum, $A_{LW} \to I_E$.

Based on (16), the proportionality γ-factor in (7) is calculated, $\gamma = 1.196235$. Meanwhile, the β-factor for UCAR, 0.8354, is obtained simultaneously, which is also the maximum value for the β-factor, as shown in Fig. 3. Furthermore, the calculations are made for the surface emissivity $\varepsilon_E < 1$. Note that the proportionality γ-factor is kept as a constant once it is determined in the first place, whilst neither additional parameters nor approximation are applied.

4 Discussion

4.1 Connecting radiation to cumulative absorption

In line with Kirchhoff's law, absorptivity and emissivity is often considered as identical in a non-scattering optical medium. In the case of the atmosphere, this implies that the absorbed radiation energy in each thin layer of an atmospheric model is completely emitted in the form of photons without being transformed into internal thermal energy in the atmosphere. Nevertheless, such an radiative transfer description seems invalid for the real atmospheric radiation where photon-particle scattering and radiation heating cannot be described by using Schwarzschild's equation. Hence, Kirchhoff's law is modified in this study with quantitative agreement with the latest observations.

In history, atmospheric radiation detected near the surface was described by using Stefan-Boltzmann law, such as the empirical equation used by Ångström [6],

$$I_A^{\downarrow} = \varepsilon_A \sigma T_A^4 \,, \tag{24}$$

where ε_A is the atmospheric emissivity, T_A is the air temperature near the surface. As the atmosphere can hardly be treated as a single isothermic layer, ε_A is in fact a random variable. Hence (24) is unfit for formulating atmospheric radiation. It has been recently shown that the atmospheric emissivity ε_A be equal to LW absorptivity a_{LW} only in the absence of clouds, see (78) in [7], but the fundamental link between atmospheric radiation and atmospheric absorption seems obscure. In general, it would appear that the distinction between the spontaneous resonant emission from the water vapor and other LW-radiation absorbers, such as CO_2, and the continuum thermal radiation governed by Planck's law remains to be further explored.

To circumvent such theoretical uncertainties, the fraction of upward *cumulative* atmospheric radiation (UCAR) at TOA, β, is introduced as a new variable in (7). In effect, the proportionality γ-factor is phenomenologically used to link the thermal radiation by the atmosphere to the *cumulative* LW atmospheric absorption based on (7). In this way, LW atmospheric radiation can be formulated. Further, the γ-factor in (7) is theoretically determined as one of the simultaneous solutions, $\gamma = 1.196235$, which appears an intrinsic invariant for the surface-atmosphere system.

4.2 Realization of the global energy balance

Because numerous energy fluxes exist between the Earth's surface and the atmosphere, it seems unlikely to identify and account all of them with absolute uncertainty. In fact, inconsistencies often arise when these different components are brought together to the power balance equation [3]. Specifically, efforts have been made to determine the net LW surface radiation, defined as the difference between the upward and downward radiation intensities,

$$I_N^{(LW)} = I_A^{\downarrow} - I_E \,. \tag{25}$$

Using the optimal estimates for $I_E = 398\,\mathrm{W\,m^{-2}}$ and $I_A^{\downarrow} = 342\,\mathrm{W\,m^{-2}}$, (25) gives $I_N^{(LW)} = -56\,\mathrm{W\,m^{-2}}$, whereas a wide range for the net LW surface radiation , $-49 > I_N^{(LW)} > -65$ $\mathrm{W\,m^{-2}}$, was predicted by individual CMIP5 models [2]. Using the net SW downward radiation, $I_{SW}^{\downarrow}(0) = 161\,\mathrm{W\,m^{-2}}$, the global mean surface net radiation,

$$I_N = I_N^{(LW)} + I_{SW}^{\downarrow}(0) \,, \tag{26}$$

is used to obtain $I_N = 105\,\mathrm{W\,m^{-2}}$, which happens to be about half way between two uncompromising values, $113\,\mathrm{W\,m^{-2}}$

τ_{LW}	a_{LW}	β	var.($\mathrm{W\,m^{-2}}$)	Zhong	[7]	[8]
0.1	0.899		OLR	–	239.4	239
0.1	1.0	a_{LW}/γ	OLR	239.4		
0.1	0.899	0.39	$I_{LW}^{(N)}$	–	64.4	
any	any	a_{LW}/γ	$I_{LW}^{(N)}$	0		
0.1	0.899		$A_{LW} + A_{SW}$	–		521
0.1	0.899	0.38	$A_{LW} + A_{SW}$	–	521.8	
0.1	1.0	$1/\gamma$	A_{max}	474.4		
0.1	1.0	0.83	$A_{LW} + A_{SW}$	474.4		
0.24	0.899	0.744	$A_{LW} + A_{SW}$	430.8		
0.33	0.8	0.67	$A_{LW} + A_{SW}$	395.1		
0.51	0.6	0.49	$A_{LW} + A_{SW}$	315.8		
0.6	0.4	0.34	$A_{LW} + A_{SW}$	236.6		
0.63	0.2	0.24	$A_{LW} + A_{SW}$	157		
0.1	0.899	–	UCAR $I_A^{\uparrow}$	–		199
0.1	0.899	0.39	UCAR $I_A^{\uparrow}$	–	199.4	
0.1	1.0	0.38	UCAR $I_A^{\uparrow}$	199.4		
0.33	0.8	0.67	UCAR $I_A^{\uparrow}$	106.8		
0.51	0.6	0.49	UCAR $I_A^{\uparrow}$	40.3		
0.6	0.4	0.34	UCAR $I_A^{\uparrow}$	3.8		
0.63	0.2	0.24	UCAR $I_A^{\uparrow}$	-13		
		$1-\beta$				
0.1	1.0	0.17	DCAR $I_A^{\downarrow}$	39.3		
0.1	0.899	–	DCAR $I_A^{\downarrow}$	–		333
0.1	0.826	0.62	DCAR $I_A^{\downarrow}$	–	332	
0.23	0.899	0.35	DCAR $I_A^{\downarrow}$	49.4		
0.33	0.8	0.33	DCAR $I_A^{\downarrow}$	52.9		
0.51	0.6	0.51	DCAR $I_A^{\downarrow}$	40.1		
0.6	0.4	0.66	DCAR $I_A^{\downarrow}$	0.77		

Table 2: Calculated thermodynamic variables (var.).

and 98 $\mathrm{W\,m^{-2}}$, estimated by Stephens *et al* [3] and Trenberth *et al* [8], respectively. To explain the remaining imbalance, both the global mean sensible heat flux and the latent heat flux were considered, knowing the lack of adequate information from direct observations. Thus, it was recommended that the surface budget estimates not be used as references [2, 8].

By introducing the net downward energy flux (NDEH) at the surface, nevertheless, such statistical estimates become unnecessary. Moreover, a number of climate scenarios previously unconsidered have been quantitatively predicted, under the same Global Energy Balance condition with zero net surface energy flux, as shown in the fourth row in Table 3. In essence, any actual thermal energy transfer between the surface and the atmosphere that appears either undefined or difficult to be measured can be implicitly treated as part of N. Note that (10) implies the net downward energy flux N_0 should be solely determined by the mean surface temperature and the surface emissivity as $I_E = \varepsilon_E \sigma T_E^4$, rather than by LW DCAR as previously taken for granted in other studies [2,3], although LW DCAR may well be part of N_0. In passing, NDEH at the surface is conceptually different from the net downward heat flux introduced by Gregory *et al* [11] to describe a hypothetical vertical radiative transfer process initiated at the TOA.

4.3 The stable range of atmospheric absorption

It is shown that the total atmospheric absorption be limited by the maximum external radiation, both from the Sun and the Earth's surface. To remain at the current equilibrium surface temperature, 289.16 K, it is theoretically predicted that the minimum of the total atmospheric absorption is close to 236 $\mathrm{W\,m^{-2}}$, being significantly lower than the value that has been assumed so far. In a recent study [7], for example, the total absorption by the atmosphere 521.8 $\mathrm{W\,m^{-2}}$ was assumed. This seems unlikely because the value is 46.6 $\mathrm{W\,m^{-2}}$ higher than the maximum atmospheric absorption, $I_E + A_{SW} = 396.4 + 78 = 474.4\,\mathrm{W\,m^{-2}}$. In another report [2], it was claimed that LW DCAR $I_A^{\downarrow} = 342\,\mathrm{W\,m^{-2}}$ which requires atmospheric absorption even higher than 521 $\mathrm{W\,m^{-2}}$.

It could be argued that such an unrealistically high atmospheric absorption is merely fabricated for invoking an imaginary greenhouse effect, bearing in mind that the average solar radiation at the TOA is 342 $\mathrm{W\,m^{-2}}$. Moreover, it is revealed that (8) and (9) are incorrect in studying the earth-atmosphere system due to the limitation associated with Kirchhoff's law in formulating thermal radiation. From those radiation and energy budget diagrams, e.g. [7–9], it is clear that (8) was used to obtain the LW atmospheric absorption, 356 $\mathrm{W\,m^{-2}}$, based on that the assumed transmitted surface radiation at the TOA is 40 $\mathrm{W\,m^{-2}}$, which yields the LW atmospheric absorptivity and the LW atmospheric transmittance equal to 89.91% and 10.01%, respectively.

By using (19), by way of contrast, the predicted LW atmospheric transmittance is close to 0.24 given the LW atmospheric absorptivity is 89.91%, as shown in Table 2, in order to satisfy the power balance condition, determined by (11). As a result, the sum of the LW and SW atmospheric absorption is 430.4 $\mathrm{W\,m^{-2}}$, instead of 521.8 $\mathrm{W\,m^{-2}}$ as previously assumed in [7,8].

Further, it is shown that the proposed theory is self-consistent as the calculated OLR at TOA from (13) is indeed a constant, independent of the LW atmospheric absorption, as indicated in (2). This implies that a previous calculation of radiation forcing by assuming a change in OLR due to CO_2-doubling [7] appears inconsistent with the definition of OLR in (1). In essence, any increase in LW atmospheric absorption will spontaneously increase in UCAR to exactly keep OLR a constant, as shown in Fig. 4, consistent with Le Chatelier's principle of thermodynamics.

4.4 Characterization of atmospheric radiation

It is found that the fraction for UCAR, β, is always larger than the portion for DCAR whenever the LW atmospheric absorptivity $a_{LW} > 60\%$. This can be explained as the fact that UCAR can easily reach outer space whereas DCAR would be increasingly attenuated towards the Earth's surface. Since DCAR is treated as part of NDEF, the difference $N - I_A^{\downarrow}$ actually represents the contribution to NDEF from other thermal energy transfer processes, both radiative or non-radiative. In fact, it is found that the cumulative downward atmospheric radiation at the surface $I_A^{\downarrow}$ is about one-fourth of NDEF, which implies that DCAR would be more effectively converted into the thermal energy towards the lower-altitude atmospheric layers as it approaches towards the surface where both the air density and the air temperature are the highest, whilst the collisions are the most frequent. Hence, the relatively low range of DCAR found in this study seems consistent with the observed stable surface temperature.

It is noted that whenever LW atmospheric absorptivity decreases to a critical value, ~40%, the total atmospheric radiation, the sum of UCAR and DCAR, becomes zero, as shown in Fig. 4, which implies that no cumulative atmospheric radiation can be detected at the TOA and the surface under this condition. This can be explained in terms of total internal absorption in the atmosphere when its internal thermal energy is insufficient to maintain its equilibrium with the surface. Under this critical condition, the atmospheric radiation is completely absorbed by the atmosphere itself. This explanation is consistent with the definitions of UCAR and DCAR whose sum become negative whenever the total atmospheric absorption is less than the net downward energy flux N in (19), required for preventing the radiation cooling at the surface. Note that once the atmosphere reaches its thermal equilibrium with the surface, the surplus LW atmospheric radiation is primarily utilized by the atmosphere to cool down itself and hence increase its entropy, rather than to warm up the surface.

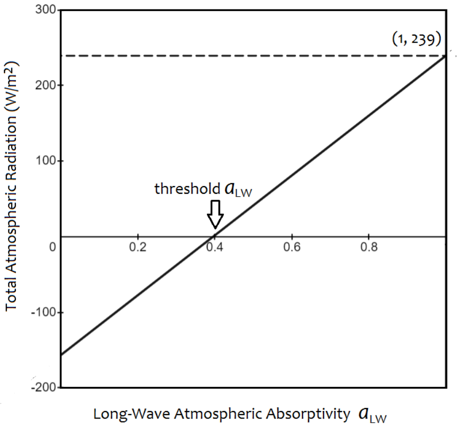


Fig. 4: Dependence of total atmospheric radiation on LW atmospheric absorptivity. Notice that the net atmospheric radiation is negative if LW atmospheric absorptivity a_{LW} is less than 0.4. The coordinate (1, 239) represents the maximum total cumulative atmospheric radiation at the TOA and the surface, 239 W m^{-2}, at the maximum LW atmospheric absorptivity a_{LW} = 100%.

ε_E	1.0	0.99	0.95	0.92	unit
I_E	396.4	392.44	376.58	364.69	W m^{-2}
$I_{LW}^{\uparrow}$	239.41	239.43	239.44	239.43	W m^{-2}
$I_{LW}^{(N)}$	0.0	0.0	0.0	0.0	W m^{-2}
N	235.4	231.44	215.58	203.69	W m^{-2}
a_{LW}^{th}	0.3971	0.391	0.3653	0.3446	
β	0.8357	0.8359	0.8354	0.8354	

Table 3: Calculated dependence on the surface emissivity.

4.5 The role of surface emissivity

The Earth's surface emissivity is explicitly treated as a thermodynamic variable in this study, whilst in the previous studies the surface emissivity was larger than 0.99 [10]. Note that the surface radiation decreases noticeably from 396.4 W m^{-2} to 364.69 W m^{-2} as the surface emissivity changes from 1.0 to 0.92 and the so-called best estimate for the surface radiance [2], 398 W m^{-2}, is 1.6 W m^{-2} higher than the calculated value at $\varepsilon_E = 1$ in this study. It is of interest to find that atmospheric radiation, both UCAR and DCAR, is independent of the surface emissivity at the maximum LW atmospheric absorptivity a_{LW} =1, although atmospheric radiation decreases non-linearly with the decrease of a_{LW}. This implies that the β-factor in (7) belongs to the intrinsic compositional properties of the atmosphere and hence independent of the intensity of the surface radiation. It is also found that LW atmospheric transmittance increases noticeably as surface emissivity changes from 1.0 to 0.92, as shown in Fig. 1, corresponding to the equilibrium NDEF decreases from 235.4 W m^{-2} to 203.69 W m^{-2}, as shown in Table 3. This indicates the atmosphere can spontaneously adjust its LW transmittance in response to the change in the surface radiance. However, such an spontaneous capability seems incapable of fully maintaining the transmitted surface radiation in the range $a_{LW} < 0.4$ unless atmospheric radiation completely ceases below each threshold value of a_{LW} for a given surface emissivity. Such detailed effects seem unexpected because the surface emissivity was often assumed as unity after Houghton [12]. Thus, the LW surface reflectivity, $r_{LW} = 1 - \varepsilon_E$, can be treated as a key variable in climate modeling. Further studies in this direction are certainly worthwhile.

5 Conclusion

In conclusion, it is shown that Kirchhoff's law on thermal radiation is oversimplified and must be modified in connecting atmospheric radiation with atmospheric absorption. Due to complicated thermal mixing processes associated with atmospheric absorption and emission, the equation for atmospheric transmittance and the atmospheric absorptivity is far from linear. Further, it is revealed that the long-wave atmospheric radiation can be completely absorbed by the atmosphere itself before it reaches to a thermal equilibrium between the surface. For the first time, both the upward cumulative atmospheric radiation and the downward cumulative atmospheric radiation can be theoretically calculated without uncertainty. It is also shown that upward cumulative atmospheric radiation at the top of the atmosphere is in general stronger than downward cumulative atmospheric radiation at the Earth's surface. It is explained that the atmospheric absorption only plays a passive role in achieving its thermal equilibrium with the Earth's surface whilst atmospheric radiation plays a proactive role in enabling the atmosphere to adapt to a wide range of variation in the atmospheric absorptivity values. In essence, only a small fraction of the atmospheric radiation, less than 55 W m^{-2}, can be absorbed by the surface, whereas the larger portion of the atmospheric radiation, up to 199 W m^{-2}, can spontaneously escape into the outer space, providing a unique mechanism for radiation cooling to maximize the entropy of the atmosphere. It is shown that the Global Energy Balance can be realized in a number of climate scenarios without any estimates. It is expected that the proposed theory can be applied in elucidating commonly concerned climate issues without invoking Kirchhoff's law and the greenhouse effect.

Acknowledgements

The author thanks H. Harde from Helmut-Schmidt University for his detailed comments and W. Happer for his encouragement at the early stage of this study. All of the graphical visualization and numerical calculation are carried out at desmos.com free of charge.

Received June 10, 2021

References

1. Ohmura A. The Development and the Present Status of Energy Balance Climatology. *J. Meteorl Soc. Jpn*, 2014, v. 92 (4), 245–285.
2. Wild M. Progress and challenges in the estimation of the global energy balance. *AIP Conference Proceedings*, 2017, v. 1810, 020004.
3. Stephens G. L., Li J., Wild M., Clayson C. A., Loeb N., Kato S., L'Ecuyer T., Stackhouse Jr. P. W., Lebsock M., and Andrews T. An Update on Earth's energy balance in light of the latest global observations. *Nature Geoscience*, 2012, v. 5 (10), 691–696.
4. Robitaille P. M. and Crothers S. J. The Theory of Heat Radiation Revisited: A Commentary on the Validity of Kirchhoff's Law of Thermal Emission and Max Planck's Claim of Universality. *Progress in Physics*, 2015, v. 11 (2), 120–132.
5. Zhong Y. C., Pilbrow J. R. A Consistent Description of EPR Spectra. *J. Magn. Reson.*, 1991, v. 93, 447–457.
6. Ångström A. Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation. *Q. J. R. Meteorol. Soc.*, 1924, v. 50 (210), 121–126.
7. Harde H. Advanced Two−Layer Climate Model for the Assessment of Global Warming by CO^2. *Open J. Atmos. Climate Change*, 2014, 2374–3794. ISSN(Online), 2374–3808.
8. Trenberth J. E., Fasullo J. T., and Kiehl J. Earths Global Energy Budget. *Bulletin of the American Meteorological Society*, 2009, v. 90 (3), 311–323.
9. Kiehl J. T., and Trenberth K. E. Earth's annual global mean energy budget. *Bull. Am. Meteorol. Soc.*, 1997, v. 78 (2), 197–208.
10. Wilber A. C., Kratz D. P., and Gupta S. K. Surface emissivity maps for use in satellite retrievals of longwave radiation. *NASA Tech. Publication*, 1999, NASA/TP-1999-209362, 35 pp.
11. Gregory J. M., Ingram W. J., Palmer M. A., Jones G. S., Stott P. A., Thorpe R. B., Lowe J. A., Johns T. C., and Williams K. D. A new method for diagnosing radiative forcing and climate sensitivity. *Geophys. Res. Lett.*, 2004, v. 31, L03205.
12. Houghton J. T. The Emissivity of the Earth's Surface. *Nature*, 1959, v. 183, 367.

Explicit Values for Gravitational and Hubble Constants from Cosmological Entropy Bound and Alpha-Quantization of Particle Masses

Alexander Kritov

E-mail: alex@kritov.ru, Moscow, Russian Federation.

In this study, we propose a method to derive expressions and numerical values for the gravitational and Hubble constants employing the "reverse engineering" path approach. Using the explicit form of Bekenstein bound for the cosmological horizon, we show that Nambu's mass-formula (the empirical alpha-quantization of elementary particle masses) is related to the proposed 3-D analogy of the holographic principle. The obtained form in the "median case" leads to the expression for the Hubble constant with the value of $H_0 = 71.995$ km/s/Mpc. The accuracy of obtained H_0 allows precise numerical calculation of the cosmological entropy bound, which coincides with the number of $\left(216 \times 2^{128}\right)^3$ bits. Conversely, the number leads to the expression for the gravitational constant resulting in $G = 6.67437305 \times 10^{-11}$, which exactly fits into the CODATA2018 value and the AAFII(2018) measurement [32]. As a coincidence, the proposed approach combined with the previous formulation of the LNNC (Large Numerical Number Coincidences) [10] allows obtaining the numerical value for the proton-to-electron mass ratio μ with an accuracy of 10^{-6}.

Note: SI units are deployed.

1 Introduction

Dimensionless numbers, including the mass ratios of the elementary particles and large numbers introduced by Dirac [45], remain unresolved puzzles in physics. To understand the significance of large number relations, the constant H_0 and G must be precisely measured, and their deviations in time and space must be constrained. However, at present, we know the value of the Hubble constant with a precision of $< 10\%$ [35]. Today, those using Planck and cosmic background data to measure a value for the Hubble constant arrive at a figure of 67.4 ± 0.5. However, the local approach provides a figure of 73.5 ± 2 [41, 43]. In contrast, the gravitational constant G is known to have better precision; however, its value has a relative accuracy of 2×10^{-5} depending on the measurement methods performed.

This paper presents an attempt to connect the Bekenstein cosmological entropy bound with the alpha-quantization of elementary particle masses. As a result, the Dirac large numbers appear as an intrinsic property of the cosmological entropy bound, which allows us to obtain the numerical value and to express G and H_0.

In 1952, Nambu proposed an empirical formula for the mass spectrum of elementary particles, known as "alpha quantization" [1]

$$m_n \approx \frac{n}{2}\, \alpha^{-1}\, m_e \tag{1}$$

where n is a natural number, α is the fine structure constant, and m_e is the electron mass. The mass interval is predicted as a half-integral multiple of approximately 70 MeV. It provides the muon mass with $n = 3$, the pion mass for $n = 4$, and the proton mass for $n = 27$ *etc*. Currently, at least 21 fundamental particles with lifetimes $>10^{-24}$ s are covered by this rule, with deviations of less than 1% [9]. The alpha quantization of elementary particle masses is extensively reviewed in the modern literature [16–28]. In particular, it is valid, for example, for the heaviest known particle, the top quark for which $n = 137 \times 36$ [20]. The Nambu formula was derived empirically and did not have any theoretical background. However, along with the new approaches to elaborate it in the frame of modern models, there were a few almost forgotten attempts to refine the formula, for example, by Nambu in 1966 [2], Hermann [3], and later [36–39] extending the quantum oscillator model, which led to clarifications and more accurate results for the mass ratios of elementary particles.

2 Bekenstein entropy bound for cosmological horizon

The cosmological (Hubble or de Sitter) horizon corresponds to the radius and volume.

$$R_H = \frac{c}{H_0}\,, \quad V_H = \frac{4\pi}{3}\left(\frac{c}{H_0}\right)^3, \tag{2}$$

where H_0 denotes the Hubble constant. Because we are looking for the upper limit of entropy, we shall consider the entire mass-energy content of the universe with $\Omega_{Tot} = 1$. Therefore, the critical density $\rho_{cr} = 3H_0^2/8\pi G$ within the Hubble volume provides the mass-energy

$$E = V_H\, \rho_{cr}\, c^2 = \frac{c^5}{2GH_0}\,. \tag{3}$$

It is easy to see that in such a case (i.e. $\Omega_{Tot} = 1$), the cosmological horizon also coincides with the Schwarzschild black hole radius*. The Bekenstein entropy bound for the black hole is

$$S = \frac{\lambda R E}{\hbar c} = \frac{4\pi R E}{\hbar c}. \tag{4}$$

The original Bekenstein formula [30] was derived based on the general considerations for "an arbitrary system of effective radius R" and contains factor $\lambda = 2\pi$. Recently the factor was clarified [31]; it was explicitly shown that particularly in the application to the Schwarzschild black hole case the factor is $\lambda = 4\pi$, which is strictly derived based on the entropy associated with the Hartle-Hawking state. Since the cosmological horizon coincides with the Schwarzschild black hole radius, as shown, the expression (4) has $\lambda = 4\pi$. The substitution of R and E from (3) leads to the value of the upper bound for the entropy of the universe

$$S = 2\pi \frac{c^5}{\hbar G H_0^2}. \tag{5}$$

The number (measured as the number of bits†) is known by its order and is also referred to as the computational capacity of the universe [44]. Notably, the critical mass of the universe can be written in terms of the obtained expression for the total entropy:

$$M_U = \frac{c^3}{2\,G H_0} = S \times m_0\,, \quad m_0 := \frac{\hbar H_0}{4\pi c^2}. \tag{6}$$

Hence, the mass m_0 can be interpreted as the minimal possible quanta of the mass-energy carrying one bit of information.

Note on Oldershaw-Fedosin scaling of the Planck constant

Using (5), one may consider the "scaled" Planck constant $\hbar^*$ such that‡

$$\hbar^* = \frac{S}{2\pi}\,\hbar, \quad h^* = S\,\hbar. \tag{7}$$

The constant $\hbar^*$ plays the role of the reduced Planck constant in a multiverse, where our universe represents an elementary particle or a quantum oscillation [4,5]. The Heisenberg uncertainty relation, which is hypothetically valid in a multiverse, is then given as

$$\Delta E\,\Delta\!\left(\frac{1}{H_0}\right) \geq \frac{\hbar^*}{2}. \tag{8}$$

On the other hand, the substitution of (7) into the expression leads to the Bekenstein law, which bounds the entropy by corresponding the total energy and time $1/H_0$ (or radius c/H_0) for the universe. Such notable correspondence to the Heisenberg uncertainty relation for the cosmological case is possible when using the above-mentioned factor $\lambda = 4\pi$ for the cosmological case.

3 3-D analogue of Holographic Principle with the "cell of space volume"

The Bekenstein bound implies the holographic principle [29]. Applying it to the cosmological horizon, the Hubble area can be represented as

$$A_H = 4\pi \left(\frac{c}{H_0}\right)^2 = S \times A_{Pl}\,, \tag{9}$$

where A_{Pl} is the Planck area§ and S is the Bekenstein cosmological bound (5). As the Plank area plays the role of a 1-bit unit of the area, analogous to that we may also introduce "the cell of space volume" V_0 such that the total Hubble volume is

$$V_H = \frac{4\pi}{3}\left(\frac{c}{H_0}\right)^3 = S \times V_0\,. \tag{10}$$

Thus, the introduced V_0 shall play a similar role for 3-D being a 1-bit unit for the volume, as the Planck area does for 2-D. The substitution of (5) leads to the explicit form

$$V_0 = \frac{V_H}{S} = \frac{2}{3}\,\frac{G\hbar}{H_0 c^2}\,. \tag{11}$$

This parameter V_0 was introduced in the author's previous work [12, 13]. The new parameter of the space volume cell V_0 may imply a different sense than *the grain of space* used in the loop quantum gravity (LQG) approach [46, 47], where the grain of space is considered to be of the order of Planck length l_{Pl}^3. In contrast, the volume cell V_0 is of the order of the cube of the reduced Compton wavelength of an elementary particle. Simultaneously, similar constraints are given for the V_0-dependent uncertainty relation in the LQG approach [47].

4 The Nambu formula for alpha-quantization of particle masses

The V_0-dependent uncertainty relation is:

$$\left(\frac{1}{2}\,\frac{\hbar}{mc}\right)^3 \geq V_0\,. \tag{12}$$

Based on that, one may consider the quantization of elementary particle masses (1) as a classic quantum harmonic oscillator [40]. The particles' rest masses correspond to the oscillator eigenstates

$$E_n = m_n c^2 = \frac{n}{2}\,\hbar\omega\,, \quad \omega = \frac{c}{L}\,,$$

where $L = V_0^{1/3}$ is the characteristic length of the oscillator, and n is a natural number for both parity cases with non-zero

---

[*] Since $R_S = (2G/c^2)(c^3/2GH_0)$.

[†] The entropy S is the number of states, the exact number of the Plank areas in covering area when using the holographic principle (9). Hence, factor ln(2) in the Bekenstein expression to obtain the number of bits, which appears in many textbooks is highly arguable and shall not be used.

[‡] Here it would be natural to introduce the "reduced" $\bar{S} = S/2\pi$ such that (7) takes the simpler form $\hbar^* = \bar{S}\hbar$, $h^* = \bar{S}\,h$.

[§] In such a way, the Plank area acquires a prefactor of two as $A_{Pl} = 2G\hbar/c^3$.

ground state ($n = 1$). Therefore, for elementary particles with mass m_n, the following condition holds:

$$\frac{n}{2}\,\frac{\hbar}{m_n c} = V_0^{1/3}\,. \tag{13}$$

The substitution of (11) for V_0 leads to particle masses that satisfy the above requirement

$$m_n = \frac{n}{2}\left(\frac{3}{2}\,\frac{\hbar^2 H_0}{Gc}\right)^{1/3}\,. \tag{14}$$

By direct calculation, it can be noted that

$$\left(\frac{3}{2}\,\frac{\hbar^2 H_0}{Gc}\right)^{1/3} \approx 137\, m_e\,,$$

where m_e is the mass of an electron. Thus, the obtained expression (14) represents Nambu's original mass formula (1), which is now related to the Bekenstein cosmological bound. The exact match to the factor to α^{-1} is achieved when the Hubble constant is $H_0 = 71.9949$ km/s/Mpc, as reviewed in the next section.

5 The Hubble constant, the Universe entropy number and *G* in the "median" case

Considering the "median case" or the "ideal" case when the exact equality in (14) holds as

$$\left(\frac{3}{2}\,\frac{\hbar^2 H_0}{Gc}\right)^{1/3} = \alpha^{-1}\, m_e\,, \tag{15}$$

it becomes possible to express the Hubble constant via better known G as

$$H_0 = \frac{2}{3}\,\frac{G\,c\,\alpha^{-3}\,m_e^3}{\hbar^2} = \frac{2}{3\alpha}\,\frac{Gm_e}{r_e^2\,c}\,, \quad \text{where } r_e = \frac{ke^2}{m_e c^2}\,, \tag{16}$$

which results in 71.9949 km/s/Mpc or 2.333×10^{-18} s^{-1} when using CODATA2018 for G. Substituting H_0 into (11) yields

$$V_0 = \left(\frac{ke^2}{m_e c^2}\right)^3 = r_e^3\,. \tag{17}$$

Furthermore, the substitution of the obtained H_0 (16) into (5) yields an explicit value for the universe total entropy bound:

$$S = \frac{4\pi}{3}\left(\frac{ke^2}{Gm_e^2}\,\frac{3\alpha}{2}\right)^3\,. \tag{18}$$

The obtained expression allows the accurate calculation of the value as 3.9711×10^{122} till the 5th digit (corresponding to the accuracy of G). Moreover, because we expect the entropy S to be a *natural number* (number of bits of information), and as binary, it most probably should contain powers of 2. The search leads to the number that represents the cosmological entropy bound as a factor of two first primes

$$S = 3^9 \times 2^{393} = \left(216\times 2^{128}\right)^3\,. \tag{19}$$

Remarkably, the found number appears to be the cube of a natural number. The number provides a sufficient relative accuracy of 3×10^{-5} with (18) corresponding to the accuracy of G (see Section 8 for a more detailed discussion on this number). Furthermore, the reverse substitution of the number to (18) allows us to express the gravitational constant:

$$G = \frac{ke^2}{m_e^2}\left(\frac{3\alpha}{2}\right)\left(\frac{4\pi}{3S}\right)^{1/3} = \left(\frac{4\pi}{3}\right)^{1/3}\frac{\alpha}{144}\,\frac{ke^2}{m_e^2}\,2^{-128} \tag{20}$$

resulting in $G = 6.67437305\times10^{-11}$. This value perfectly fits the value of CODATA2018 for G. The obtained value also coincides with the AAFII(2018) measurement of 6.674375(82) performed with very high precision [32]. Moreover, the use of the obtained G in (16) results in the expression for the Hubble constant

$$H_0 = \frac{c}{r_e}\left[\left(\frac{4\pi}{3}\right)^{1/3}\frac{1}{216}\right]2^{-128}\,, \tag{21}$$

where r_e is the classical electron radius (16). Notably, to satisfy the equality to α in (15), the expressions acquire the factor given in square brackets. Denoting this factor as $\alpha_s = 1/133.995..$ (or "alpha-substitute"), both expressions can be written in the simpler form

$$H_0 = \alpha_s\,\frac{c}{r_e}\,2^{-128}\,, \quad G = \alpha_s\,\alpha\,\frac{3}{2}\,\frac{ke^2}{m_e^2}\,2^{-128}\,, \tag{22}$$

where

$$\alpha_s := \left(\frac{4\pi}{3}\right)^{1/3}\frac{1}{216}\,.$$

The significance of this parameter is reviewed further.

6 Proton to electron mass ratio from deviated *G* and H_0

We have considered the "median" or ideal case of exact equality to α in the extended Nambu's mass formula (14). In a real-life scenario, the masses of the elementary particles deviate from the median values by ±1%. There are two alternative ways to refine the Nambu mass formula to obtain more accurate masses for elementary particles. The first approach, as mentioned in Section 1, clarifies the quantum oscillator model. This leads to the appearance of eigenvalues or zeros of some functions instead of the natural number n in (1). The second alternative is to introduce the deviation of G and H_0 in the mass formula (14), which would also lead to nonconstancy of V_0 and deviated states of the entropy from S depending on the nature of the particle. The first method appears to be preferable and requires further studies using QM.

However, in this section, we evaluate the second "heuristic" alternative.

One may recall the previously proposed expressions for G and the Hubble constants (LNNC) [10]. Denoting them with a prime (′) to distinguish them from the obtained "median" values, they are

$$H_0' = \frac{m_e c^2}{\hbar} 2^{-128}, \quad G' = \frac{3}{20} \frac{ke^2}{m_e m_p} 2^{-128}, \tag{23}$$

where m_p is the proton mass. The value for $H_0' = 70.39882$ km/s/Mpc differs by 2%* from the "median" case of H_0 (16). It is evident that the formula for H_0' can be treated as derived from the expression for H_0 (22), wherein the latter, α_s is simply *substituted* by α.

The value of the gravitational constant is $G' = 6.6746305 \times 10^{-11}$, which deviates from the median G (22) by 3×10^{-5}, and is closer to the AAF-I and AAF-III measurements [32]. It is evident that these values (23) do not provide the equality to "alpha" in mass-formula (15); however remarkably H_0' and G' being substituted into (14) with $n = 27$ provide a good approximation of the proton mass, thus the ratio becomes

$$\mu = \frac{m_p}{m_e} = \left(\frac{5}{4} 3^9 \alpha^{-1}\right)^{1/2}. \tag{24}$$

Moreover, it can be seen that both suggested formulas for the gravitational constant have relative deviations of 3×10^{-5}, equating G from (23) and (22) gives

$$\alpha^{-1} = \left(\frac{4\pi}{3}\right)^{1/3} \frac{5\mu}{108}, \tag{25}$$

where we expect the same relative error of 10^{-5}. The substitution of μ from (24) leads to

$$\alpha^{-1} = \left(\frac{4\pi}{3}\right)^{2/3} \left(\frac{15}{4}\right)^3 = \mathbf{137.03}12258, \tag{26}$$

and substituting it again to (24) results in

$$\mu = \left(\frac{4\pi}{3}\right)^{1/3} \left(\frac{135}{4}\right)^2 = \mathbf{1836.15}959, \tag{27}$$

which has a relative accuracy with the experimental value of the proton-to-electron ratio of 3×10^{-6}. The remarkable property of both expressions is their simple forms that involve powers of the first three primes as $15 = 5 \times 3$ and $135 = 3^3 \times 5$. The expression for μ can be assumed to be the best in terms of the precision-simplicity ratio (see [11] to see the complication level of formulae with comparable accuracy for μ). The expressions can also be rewritten in the following forms:

$$\mu = \frac{27}{2} (135\,\phi), \quad \alpha^{-1} = 135\,\phi^2, \tag{28}$$

where

$$\phi := \frac{5}{8} \left(\frac{4\pi}{3}\right)^{1/3},$$

and the "alpha-substitute" is explicitly $\alpha_s^{-1} = 135\,\phi^{-1}$. Thus, the formula for μ restores the original form of Nambu's mass formula with $n = 27$. Hence, the factor $\phi \approx 1.0075$ plays the role of a small deviation and exhibits a deviation of α in Nambu's mass formula. Simultaneously, ϕ^2 shows how α deviates from an integer of 135. The deviation $\phi^3 \approx 1.02$ also provides the explicit ratios of the two values for the Hubble constants given by the expressions (21) and (23) as this deviation is given by the ratio of "alpha-substitute" in (21) to the exact "alpha" in (23).

7 Quantum number of the Universe and Eddington's number of particles

The paper would not be complete without reviewing the Eddington number of particles (pairs of protons and electrons), which he assumed to be $N = 2 \times 136 \times 2^{256}$ [6]. In Section 5, we review the number for the Bekenstein entropy S, which is also expressed by the power of 2 (19). Prior to the calculation of the Eddington number of particles, we calculate the n-number using the obtained mass formula (14) applied to the *entire* universe mass with $\Omega = 1$

$$M_u = \frac{c^3}{2GH_0} = \frac{n_u}{2} \left(\frac{3}{2} \frac{\hbar^2 H_0}{Gc}\right)^{1/3}. \tag{29}$$

Because we are applying it to "the median" case, it is clear that $M_u = (n_u/2)(\alpha^{-1} m_e)$. Using the obtained values for G and H_0 (22) after a few manipulations, the number becomes

$$n_u = \frac{2}{3} \left(\frac{3}{4\pi} S\right)^{2/3} = \frac{2}{3} \alpha_s^{-2}\, 2^{256}. \tag{30}$$

Using this number, it is evident that the Eddington number of protons can be expressed as

$$N = \frac{M_u \Omega_{M_p}}{m_p} = \Omega_{M_p} \frac{n_u}{2} \frac{\alpha^{-1}}{\mu}, \tag{31}$$

where Ω_{M_p} is the proton content of the universe. The obtained good approximations for α and μ of (28) provide the ratio

$$\frac{\alpha^{-1}}{\mu} = \frac{2}{27} \phi,$$

and substituting n from (30) results in the number of protons in the universe

$$N = \Omega_{M_p} \frac{10}{3} \alpha_s^{-1}\, 2^{256}, \tag{32}$$

where the second power of α_s^{-1} decreases with ϕ. This expression is fairly close to the famous Eddington number. However, the difference is that it contains the prefactor, and "alpha-substitute" (≈ 134) instead of 136 in Eddington's expression.

---

[*] The ratio for the deviation is $(\alpha^{-2}/10\mu) = (4\pi/3)((5/8)^3)$, as can be seen later.

8 Discussion

In Section 5, we proposed the numerical value for the "median" cosmological entropy bound as the number of bits, which explicitly equals $(216 \times 2^{128})^3$. The number limits the upper bound of the informational capacity of the universe according to the Bekenstein law. When searching for the numbers of order 10^{122} with a relative accuracy of 10^{-5}, one must observe that there exist 10^{117} alternative natural numbers to choose between. Another good fit, for example, can be given by

$$137 \times \frac{81!}{2} \approx 3.971031 \times 10^{122}\,\text{bits}\,.$$

Notably, the number represents the order of the alternating group $A(81)$ with prefactor 137, which can be considered as a coincidence. However, the key advantage of (19) compared to other alternatives is that it is the simplest number composed of the product of only the first two primes. Second, it represents the cube of another natural number, which reveals its significance during the calculations. Moreover, the number (19) can also be represented by the Mersenne prime $M_{127} = 2^{127} - 1$, where M_{127} has the unique property of being the double Mersenne prime and fourth Catalan-Mersenne number* discovered by Catalan [34]. Hence, $S = (432 \times M_{127})^3$, which possibly connects the median entropy to the cyclic group† $\mathbb{Z}_{M_{127}}$.

Despite the presence of a power of two, the proposed number differs from the Eddington E-numbers [7, 8]. However, further study is required for a possible connection of the proposed number to the Clifford algebras and the finite groups of Lie type [48].

9 Conclusions

In 1935, Heisenberg [42] suggested using the number 432 to calculate the fine structure constant as $\alpha^{-1} = 432/\pi$. The paper has demonstrated that number 432 and its derivatives (108, 216) appear in the "median" or symmetric case of universe entropy bound, and further in the calculation for the dimensionless numbers (25), (19). An intriguing numerical expression for the total universe entropy for the Bekenstein cosmological bound is proposed (19), which contains only powers of 2 and 3. This allowed to construct an expression for the gravitational constant that gives a value of $G = 6.67437294 \times 10^{-11}\,\text{m}^3\,\text{kg}^{-1}\,\text{s}^{-2}$, which fits the range of CODATA2018 to the latest measurements. Along with the previous formulation for the Hubble constant, the approach provides a new alternative form (greater by 2% from the previous) as given by expressions (21), (16), resulting in $H_0 = 71.994$ km/s/Mpc, which corresponds to the "median" case of the universe entropy bound (19). The current accuracy of measurements of the Hubble constant H_0 limits the study. Further improvements in the measurements of the Hubble constant are required, as it will clarify the concordance of the value of the cosmological entropy bound S to the proposed number.

The paper proposes a path, using the explicit value for Bekenstein bound, to connect the maximum of the observable values such as the Hubble volume, the total mass of the universe with minimal measurable values (V_0 and m_0), which supposedly have to play a role in quantum mechanics. The approach can be extended towards a broader range of physical parameters, such as maximal and minimal force, maximal and minimal acceleration. The introduction of such parameters may lead to new approaches in quantum mechanics and cosmology. Further research is required in the frame of quantum mechanics, the LQG, which would utilize the introduced space volume V_0 parameter in connection to Clifford algebra $Cl_{3,0}$ (APS), where it has the correspondence to volume coordinate x_{123} [14]. Such a study may further refine the quantum oscillator model of elementary particle masses using the introduced parameters.

Received on July 9, 2021

---

*Since $127 = 2^7 - 1$, $7 = 2^3 - 1$, and $3 = 2^2 - 1$.

†Interestingly; this number also nearly coincides with the order of symplectic group Sp(n, q) with $q = 2^{43}$, $n = 1$ with prefactor 108, and the same for the Chevalley group $A_n(q)$ (PSL(n, q)).

References

1. Nambu Y. An Empirical Mass Spectrum of Elementary Particles. *Prog. Theor. Phys.*, 1952, v. 7, 131.
2. Nambu Y. Relativistic Wawe Equations for Particles with Internal Structure and Mass Spectrum. *Progr. Theoret. Phys. (Kyoto)*, 1966, Suppl., Nos. 37–38, 368–382.
3. Hermann R. General Mass Formula for the Nambu Wave Equations. *Physical Review*, 1968, v. 167, 525.
4. Fedosin S. G. Physics and Philosophy of Similarity from Preones to Galaxies. Perm, 1999 [in Russian].
5. Oldershaw R. The Hidden Meaning of Planck's Constant. *Universal Journal of Physics and Application*, 2013, v. 1 (2), 88–92.
6. Eddington A. S. Relativity Theory of Protons and Electrons. Cambridge University Press, Cambridge, 1936.
7. Eddington A. S. On Sets of Anticommuting Matrices. Part II: The Factorization of E-Numbers. *Journal of the London Mathematical Society*, 1933, v. 1-8 (2), 142–152.
8. Salingaros N. Some remarks on the algebra of Eddington's E numbers. *Foundations of Physics*, 1985, v. 15, 683–691.
9. Greulich K. O. Calculation of the Masses of All Fundamental Elementary Particles with an Accuracy of Approx. 1%. *J. Mod. Phys.*, 2010, v. 1, 300–302.
10. Kritov A. A New Large Number Numerical Coincidences. *Progress in Physics*, 2013, v. 9 (2), 25–28.
11. Kritov A. An Essay on Numerology of the Proton to Electron Mass Ratio. *Progress in Physics*, 2015, v. 11 (1), 10–13.
12. Kritov A. On the Fluid Model of the Spherically Symmetric Gravitational Field. *Progress in Physics*, 2019, v. 15 (2), 101–105.
13. Kritov A. Unified Two Dimensional Spacetime for the River Model of Gravity and Cosmology. *Progress in Physics*, 2019, v. 15 (3), 163–170.
14. Kritov A. Gravitation with Cosmological Term, Expansion of the Universe as Uniform Acceleration in Clifford Coordinates. *Symmetry*, 2021, v. 13, 366.
15. PDG, Physical Review D, 2020.

16. Jensen E. Regularities in the masses of some elementary particles. RVAU-IMS-80-2, Veterin. Agricult. Univ. Dept. Math. Stat., Copenhagen, 1980.
17. Giani S. Particle Mass-Formulae. CERN-OPEN-2004-004, 02 March 2004.
18. Mac Gregor M. H. Electron generation of leptons and hadrons with reciprocal α-quantized lifetimes and masses. arXiv: hep-ph/0506033.
19. Mac Gregor M. H. The top quark to electron mass ratio. arXiv: hep-ph/0603201.
20. Mac Gregor Malcolm H. A "Muon Mass Tree" with α-quantized Lepton. Quark and Hadron Masses. arXiv: hep-ph/0607233.
21. Mac Gregor M. H. Models for Particles. *Lett. Nuovo Cim.*, 1970, v. 7, 211–214.
22. Mac Gregor M. H. The Power of Alpha: The Electron Elementary Particle Generation with Alpha-Quantized Lifetimes and Masses. World Scientific Publishing, Singapore, 2007.
23. Palazzi P. Particles and Shells. CERN-OPEN-2003-006, 2003. arXiv: physics.gen-ph/0301074.
24. Palazzi P. The meson mass system. *Int. J. Mod. Phys.* 2007, v. 22, 546–549.
25. Shah G. N., Mir T. A. Pion and muon mass difference: a determining factor in elementary particle mass distribution. *Mod. Phys. Lett. A.*, 2008, v. 23, 53.
26. Mir T. A., Shah G. N. Order in the mass spectrum of elementary particles. arXiv: physics.gen-ph/0806.1130.
27. Greulich K. O. Calculation of the Masses of All Fundamental Elementary Particles with an Accuracy of Approx. 1%. *J. Mod. Phys.*, 2010, v. 1, 300–302.
28. Chiatti L. A Possible Model for the Mass Spectrum of Elementary Particles. *Phys. Essays.*, 2012, v. 25, 374–386.
29. Bekenstein J. D. Information in the Holographic Universe. *Scientific American*, 2003, v. 289 (2), 58–65.
30. Bekenstein J. D. Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D*, 1981, v. 23 (2), 287–298.
31. Longo R., Xu F. Comment on the Bekenstein bound. arXiv: math-ph/1802.07184v1, Proposition 2.5.
32. Qing Li, Chao Xue, Jian-Ping Liu, Jun-Fei *et al* Measurements of the gravitational constant using two independent methods. *Nature*, 2018, v. 560, 582–588.
33. Merkatas C., Toman B., Possolo A., Schlamminger S. Shades of Dark Uncertainty and Consensus Value for the Newtonian Constant of Gravitation. arXiv: physics.data-an/1905.09551.
34. Catalan E. Sur la théorie des nombres premiers. Turin, 1876, p. 11. Catalan, E. Théorie des nombres. 1891, p. 376. See also the OEIS A007013.
35. Jackson N. The Hubble Constant. *Living Rev. Relativity*, 2015, v. 18.
36. Salomatov V. N. Relativistic particle in the rest frame. arXiv: physics.gen-ph/1401.6253.
37. Salomatov V. N. Helmholtz equation in relativistic quantum mechanics. *Physics Essays*, 2017, v. 30, 2.
38. Sidharth B. G. A QCD Generated Mass Spectrum. arXiv: physics.gen-ph/0309037v1.
39. Ram B., Halasa R. Meson Spectra with a harmonic-oscillator potential in the Klein-Gordon equation. *Physical Review D*, 1971, v. 19, 11.
40. Landau L. D., Lifshitz E. M. Quantum Mechanics: Non-Relativistic Theory. Course of Theoretical Physics, v. 3. Second Edition. Pergamon Press, 1965, 67–72.
41. Nandita K., Luca I., Marica B. *et al*. A new measurement of the Hubble constant using Type Ia supernovae calibrated with surface brightness fluctuations. *Astronomy and Astrophysics*, 2021, v. A72, 647.
42. Kragh H. Magic number: A partial history of the fine-structure constant. Arch. Hist. Exact Sci., 2003, v. 57, 395–431.
43. Taubenberger S., *et al*. The Hubble constant determined through an inverse distance ladder including quasar time delays and Type Ia supernovae. *Astronomy and Astrophysics* 2019, v. 628, L7.
44. Lloyd S. Computational capacity of the universe. arXiv: quant-ph/0110141.
45. Dirac P. A. M. Cosmological models and the large numbers hypothesis. *Proceedings of the Royal Society of London A*, 1974.
46. Rovelli C. Loop Quantum Gravity. arXiv: gr-qc/9710008v1.
47. Bojowald M. Critical evaluation of common claims in loop quantum cosmology. arXiv: gr-qc/2002.05703v1.
48. Solomon R. A Brief History of the Classifications of the Finite Simple Groups. *American Mathematical Society*, 2001, v. 38 (3), 315–352.

Physics of Transcendental Numbers Determines Star Distribution

Hartmut Müller

Rome, Italy.
E-mail: hm@interscalar.com

Transcendental ratios of physical quantities can provide stability in complex dynamic systems because they inhibit the occurrence of destabilizing resonance between the elements of the system. This approach leads to a fractal scalar field that affects any type of physical interaction. In this paper we verify the model claims on the frequency distribution of interstellar distances in the solar neighborhood.

Introduction

Since the beginning of the past century astronomers began to routinely measure stellar parallaxes. In 1957 this effort was formalized with the publication [1] of 915 stars within 20 pc. Various updates and extensions to larger distances produced what became the Catalogue of Nearby Stars (CNS), including 3803 stars within 25 pc [2] released in 1991. Hipparcos [3] increased the quantity and quality of the CNS content. In 1998 the CNS dataset went online and currently has 5835 entries, but it is no longer updated. The most recent update [4] of the CNS was to provide accurate coordinates taken from the Two Micron Sky Survey (2MASS) [5]. Finally, the Gaia Catalogue of Nearby Stars (GCNS) attempts to make a census of all stars in the solar neighborhood using the Gaia results [6]. In the GCNS, the solar neighborhood is defined as a sphere with a radius of 100 pc centered on the Sun.

In this paper, we will analyze the distribution of the number of stars in the solar neighborhood as function of their mutual distances. This approach is not heliocentric and does not deal with fixed reference points at all.

Conventional models expect an exponential increase of the cumulative number of stars with the distance from a fixed reference point, such as the Sun. As shown in [6], this actually appears to be the case.

We will show that the consideration of all possible pairs of stars within a given range of interstellar distances leads to the appearance of a stable scale-invariant pattern in the frequency distribution of the number of stars as function of the distance between them. This means that there are interstellar distances preferred by the majority of stars in the solar neighborhood. Furthermore, we will derive this scale-invariant pattern from a number theoretic approach.

Methods

In [7] we have shown that the difference between rational, irrational algebraic and transcendental numbers is not only a mathematical task, but it is also an essential aspect of stability in complex dynamic systems. For instance, integer frequency ratios provide resonance interaction that can destabilize a system [8]. Actually, it is transcendental numbers that define the preferred ratios of quantities which avoid destabilizing resonance interaction [9]. In this way, transcendental ratios of quantities sustain the lasting stability of periodic processes in complex dynamic systems. With reference to the evolution of a planetary system and its stability, we may therefore expect that the ratio of any two orbital periods should finally approximate a transcendental number.

Among all transcendental numbers, Euler's number $e = 2.71828\ldots$ is unique, because its real power function e^x coincides with its own derivatives. In the consequence, Euler's number allows inhibiting resonance interaction regarding any interacting periodic processes and their derivatives. Because of this unique property of Euler's number, complex dynamic systems tend to establish relations of quantities that coincide with values of the natural exponential function e^x for integer and rational exponents x.

Therefore, we expect that periodic processes in real systems prefer frequency ratios close to Euler's number and its rational powers. Consequently, the logarithms of their frequency ratios should be close to integer $0, \pm 1, \pm 2, \ldots$ or rational values $\pm 1/2, \pm 1/3, \pm 1/4, \ldots$ In [10] we exemplified our hypothesis in particle physics, astrophysics, cosmology, geophysics, biophysics and engineering.

Based on this hypothesis, we introduced a fractal model of matter [11] as a chain system of harmonic quantum oscillators and could show the evidence of this model for all known hadrons, mesons, leptons and bosons as well. In [12] we have shown that the set of stable eigenstates in such systems is fractal and can be described by finite continued fractions:

$$\mathcal{F}_{jk} = \ln\left(\omega_{jk}/\omega_{00}\right) = \langle n_{j0}; n_{j1}, n_{j2}, \ldots, n_{jk} \rangle \tag{1}$$

where ω_{jk} is the set of angular eigenfrequencies and ω_{00} is the fundamental frequency of the set. The denominators are integer: $n_{j0}, n_{j1}, n_{j2}, \ldots, n_{jk} \in \mathrm{Z}$. The cardinality $j \in \mathrm{N}$ of the set and the number $k \in \mathrm{N}$ of layers are finite. In the canonical form, all numerators equal 1. We use angle brackets for continued fractions.

Any finite continued fraction represents a rational number [13]. Therefore, the ratios ω_{jk}/ω_{00} of eigenfrequencies are always irrational, because for rational exponents the natural exponential function is transcendental [14]. This circumstance provides for lasting stability of those eigenstates of a chain system of harmonic oscillators because it prevents resonance interaction [15] between the elements of the system.

The distribution density of stable eigenstates reaches local maxima near reciprocal integers $\pm 1/2, \pm 1/3, \pm 1/4, \ldots$ that are attractor points (fig. 1) in the fractal set $\mathcal{F}_{jk}$ of natural logarithms. Integer logarithms $0, \pm 1, \pm 2, \ldots$ represent the most stable eigenstates (main attractors).

In the case of harmonic quantum oscillators, the continued fractions $\mathcal{F}_{jk}$ define not only fractal sets of natural angular frequencies ω_{jk}, angular accelerations $a_{jk} = c \cdot \omega_{jk}$, oscillation periods $\tau_{jk} = 1/\omega_{jk}$ and wavelengths $\lambda_{jk} = c/\omega_{jk}$ of the chain system, but also fractal sets of energies $\mathrm{E}_{jk} = \hbar \cdot \omega_{jk}$ and masses $\mathrm{m}_{jk} = \mathrm{E}_{jk}/c^2$ which correspond with the eigenstates of the system. For this reason, we call the continued fraction $\mathcal{F}_{jk}$ the *Fundamental Fractal* of stable eigenstates in chain systems of harmonic quantum oscillators.

Fig. 1: The distribution of stable eigenvalues of $\mathcal{F}_{jk}$ for $k = 1$ (above) and for $k = 2$ (below) in the range $-1 \leqslant \mathcal{F}_{jk} \leqslant 1$.

The spatio-temporal projection of the Fundamental Fractal $\mathcal{F}_{jk}$ of stable eigenstates is a fractal scalar field of transcendental attractors, the *Fundamental Field* [16].

The connection between the spatial and temporal projections of the Fundamental Fractal is given by the speed of light $c = 299792458$ m/s. The constancy of c makes both projections isomorphic, so that there is no arithmetic or geometric difference. Only the units of measurement are different.

Figure 2 shows the linear 2D-projection $\exp(\mathcal{F}_{jk})$ of the first layer of the Fundamental Field

$$\mathcal{F}_{\mathrm{j}1} = \langle n_{j0}; n_{j1} \rangle = n_{j0} + \frac{1}{n_{j1}}$$

in the interval $-1 < \mathcal{F}_{\mathrm{j}1} < 1$. The upper part of figure 1 shows the same interval in the logarithmic representation. The Fundamental Field is topologically 3-dimensional, a fractal set of embedded spheric equipotential surfaces. The logarithmic potential difference defines a gradient directed to the center of the field that causes a central force of attraction. Because of the fractal logarithmic hyperbolic metric of the field, every equipotential surface is an attractor.

The Fundamental Field is of pure arithmetical origin, and there is no particular physical mechanism required as field source. It is all about transcendental ratios of frequencies [9] that inhibit destabilizing resonance. Therefore, we postulate the universality of the Fundamental Field that affects any type of physical interaction, regardless of its complexity.

In fact, scale relations in particle physics [11] and astrophysics [17] obey the same Fundamental Fractal (1), without any additional or particular settings. The proton-to-electron rest energy ratio approximates the first layer of the Fundamental Fractal that could explain their exceptional stability.

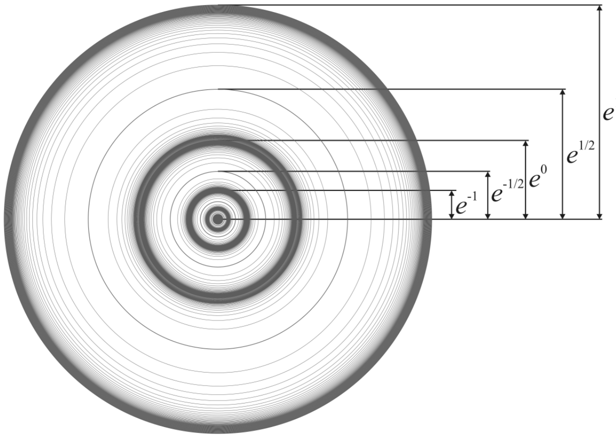


Fig. 2: The equipotential surfaces of the Fundamental Field in the linear 2D-projection for $k = 1$.

In fact, the life-spans of the proton and electron top everything that is measurable, exceeding 10^{29} years [18].

Property	Electron	Proton
$\mathrm{E} = \mathrm{mc}^2$	0.5109989461(31) MeV	938.2720813(58) MeV
$\omega = \mathrm{E}/\hbar$	$7.76344 \cdot 10^{20}$ Hz	$1.42549 \cdot 10^{24}$ Hz
$\tau = 1/\omega$	$1.28809 \cdot 10^{-21}$ s	$7.01515 \cdot 10^{-25}$ s
$\lambda = c/\omega$	$3.86159 \cdot 10^{-13}$ m	$2.10309 \cdot 10^{-16}$ m

Table 1: The basic set of the physical properties of the electron and proton. Data from Particle Data Group [18]. Frequencies, oscillation periods and wavelengths are calculated.

The proton-to-electron ratio (tab. 1) approximates the seventh power of Euler's number and its square root:

$$\ln\left(\frac{\lambda_e}{\lambda_p}\right) = \ln\left(\frac{3.86159 \cdot 10^{-13}\,\mathrm{m}}{2.10309 \cdot 10^{-16}\,\mathrm{m}}\right) \simeq 7 + \frac{1}{2} = \langle 7; 2 \rangle$$

In the consequence of this potential difference of the proton relative to the electron, the scaling factor $\sqrt{e} = 1.64872\ldots$ connects attractors of proton stability with similar attractors of electron stability in alternating sequence. The following Diophantine equation describes the correspondence of proton calibrated attractors n_p with electron calibrated attractors n_e. Non considering the signature, only three pairs (n_p, n_e) of integers are solutions to this equation: (3, 6), (4, 4), (6, 3).

$$\frac{1}{n_p} + \frac{1}{n_e} = \frac{1}{2}$$

Figure 3 demonstrates this situation on the first layer of the Fundamental Fractal (1). Both, the attractors of proton and electron stability are represented at the first layer, so we can see clearly that among the integer or half, only the attractors $\pm 1/3$, $\pm 1/4$ and $\pm 1/6$ are common. In these attractors, proton stability is supported by electron stability and vice versa, so we expect that they are preferred in real systems.

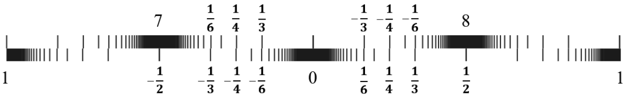


Fig. 3: The distribution of the attractors of proton (bottom) stability in the range $-1 < \mathcal{F} < 1$ of the attractors of electron (top) stability. Natural logarithmic representation.

These unique properties of the electron and proton predestinate their physical characteristics as fundamental units. Table 1 shows the basic set of electron and proton units that can be considered as a fundamental metrology (c is the speed of light in a vacuum, $\hbar$ is the Planck constant). In [12] was shown that the fundamental metrology (tab. 1) is completely compatible with Planck units [19]. Originally proposed in 1899 by Max Planck, these units are also known as natural units, because the origin of their definition comes only from properties of nature and not from any human construct. Max Planck wrote [20] that these units, "regardless of any particular bodies or substances, retain their importance for all times and for all cultures, including alien and non-human, and can therefore be called natural units of measurement". Planck units reflect the characteristics of space-time.

We assume that scale invariance according to the Fundamental Fractal (1), which is calibrated to the physical properties of the proton and the electron, is a universal characteristic of organized matter and criterion of stability. This hypothesis we have called *Global Scaling* [10].

In this paper we will show that the distribution of interstellar distances in the solar neighborhood corresponds with the distribution of attractors in the Fundamental Field.

Results

In [21] we applied the Fundamental Fractal (1) to macroscopic scales interpreting gravity as attractor effect of its stable eigenstates. Indeed, the orbital and rotational periods of planets, planetoids and large moons of the solar system correspond with attractors of electron and proton stability [12]. This is valid also for the planets [10] of the systems Trappist 1 and Kepler 20. Planetary and lunar orbits [17] correspond with equipotential surfaces of the Fundamental Field.

Figure 4 shows the distribution of the number of exoplanets with orbital periods in the range 5 d $< T <$ 24 d that corresponds with logarithms $59.2 < ln\ (T/2\pi\tau_e) < 60.8$ on the horizontal axis. According with table 1, τ_e is the electron angular oscillation period. The histogram contains data of 1430 exoplanets and shows clearly the maximum corresponding with the main attractor $\mathcal{F}\langle 60\rangle$. Other maxima correspond with the attractors $\mathcal{F}\langle 59; 2\rangle$ and $\mathcal{F}\langle 60; 2\rangle$; even the subattractors $\mathcal{F}\langle 60; -4\rangle$ and $\mathcal{F}\langle 60; 4\rangle$ can be distinguished.

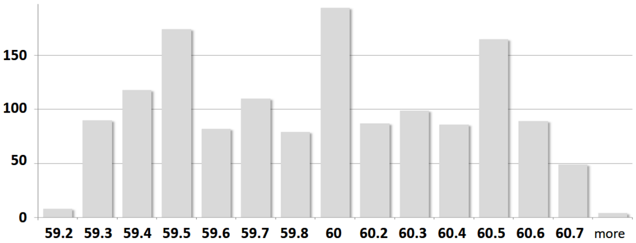


Fig. 4: The histogram shows the distribution of the number of exoplanets with orbital periods in the range 5 d $< T <$ 24 d. The logarithms ln $(T/2\pi\tau_e)$ are on the horizontal axis. Corresponding with table 1, τ_e is the electron angular oscillation period. Data of 1430 exoplanets are taken from [22].

The histogram evidences that the majority of the 1430 exoplanets [22] prefer orbital periods close to 10–11 days corresponding with the main attractor $\mathcal{F}\langle 60\rangle$, as well as periods close to 6–7 days or close to 17–18 days corresponding with the attractors $\mathcal{F}\langle 59; 2\rangle$ and $\mathcal{F}\langle 60; 2\rangle$. Because of the logarithm 7+1/2 of the proton-to-electron ratio, the attractors $\mathcal{F}\langle 59; 2\rangle$ and $\mathcal{F}\langle 60; 2\rangle$ of *electron* stability are actually the main attractors $\mathcal{F}\langle 67\rangle$ and $\mathcal{F}\langle 68\rangle$ of *proton* stability.

Figure 5a shows the distribution of the number of stars as function of their distances R from the Sun up to 25 light-years that correspond with the logarithms $ln\ (R/\lambda_e) < 68.6$ on the horizontal axis. According with table 1, λ_e is the Compton wavelength of the electron. The histogram contains 192 distances and shows clearly the maxima corresponding with the attractors $\mathcal{F}\langle 67\rangle$, $\mathcal{F}\langle 67; 2\rangle$, $\mathcal{F}\langle 68\rangle$ and $\mathcal{F}\langle 68; 2\rangle$.

Knowing the right ascension, declination and distances of two stars from the Sun, it is not difficult to calculate the distance between them. In preparation of this paper, the mutual distances between the 192 best measured stars including Vega within a radius of 25 light-years around the Sun were calculated. The number of pairs of stars is given by the formula:

$$P = N(N-1)/2$$

where N is the number of stars; P is the number of pairs. For 192 stars, we calculated $P = 18,336$ interstellar distances.

Figure 5b shows the distribution of the number of stars as function of their distances R from Sirius up to 33 light-years. Also this histogram shows clearly the maxima corresponding with the attractors $\mathcal{F}\langle 67\rangle$, $\mathcal{F}\langle 67; 2\rangle$, $\mathcal{F}\langle 68\rangle$ and $\mathcal{F}\langle 68; 2\rangle$. The same $\mathcal{F}$-pattern appears in the histograms of interstellar distances measured from Barnard's star (fig. 5b), Tau Ceti (fig. 5d) and other stars in the 25-light-years solar neighborhood. Only the expression of the $\mathcal{F}$-pattern differs in strength.

Conclusion

Standard models expecting an exponential increase of the cumulative number of stars with the distance from a fixed reference point, perhaps could interpret the local maxima in the histograms as anomalies evidencing that the solar neighborhood is still in transformation. Within our approach, on the contrary, the coincidence of the maxima with attractors of the Fundamental Field evidences that the solar neighborhood has

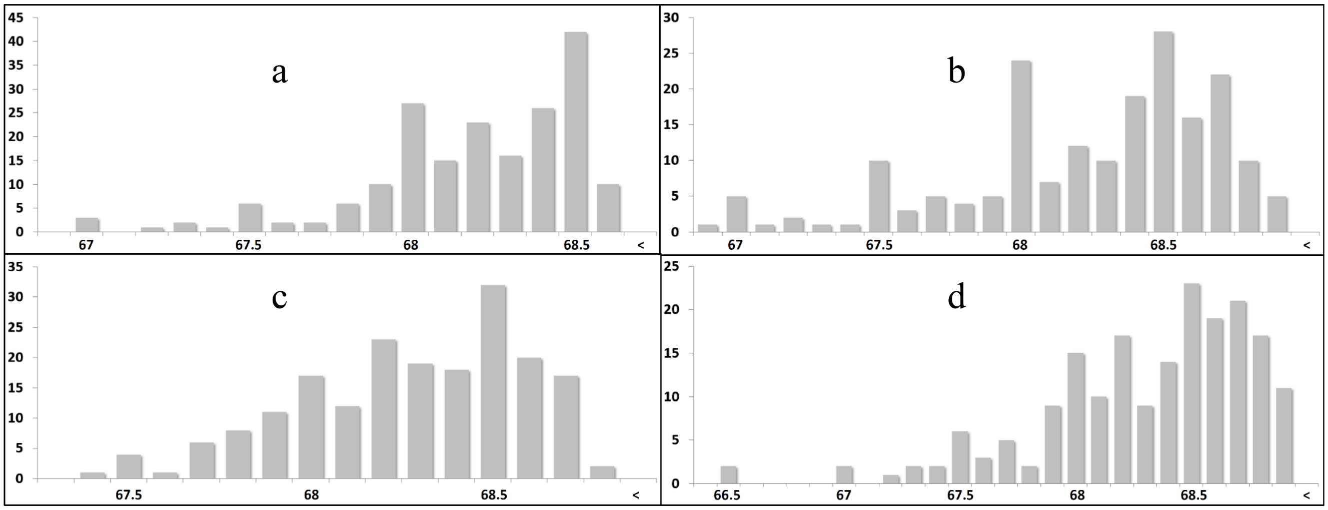


Fig. 5: The histogram shows the distribution of the number of stars in the solar neighborhood as function of their distances R from the Sun (a), Sirius (b), Barnard's star (c) and Tau Ceti (d). The logarithms ln (R/λ_e) are on the horizontal axis. Corresponding with table 1, λ_e is the Compton wavelength of the electron. Data of 192 stars are taken from [23].

already reached a certain level of stability. Moreover, we expect a continuous amplification of $\mathcal{F}$-patterns in histograms as trend of interstellar distances. Most likely, the appearance of patterns corresponding with the Fundamental Fractal (1) is a universal criterion of stability.

Since the Fundamental Fractal is of number theoretic origin, it determines the frequency distributions of interstellar distances as well as the wavelengths of elementary particles. Interscalar cosmology [10] bases on this approach.

Acknowledgements

The author is grateful to Simon Shnoll, Viktor Panchelyuga, Valery Kolombet, Oleg Kalinin, Viktor Bart, Michael Kauderer, Ulrike Granögger and Leili Khosravi for valuable discussions.

Submitted on August 27, 2021

References

1. Gliese W. Astronomisches Rechen-Institut Heidelberg Mitteilungen Serie A, 1957, v. 8, 1.
2. Gliese W., Jahreiß H. Preliminary Version of the Third Catalogue of Nearby Stars: The Astronomical Data Center CD-ROM. Selected Astronomical Catalogs, (1991).
3. Perryman M. A. C. The Hipparcos Catalogue. *Astron. Astrophys.*, 1997, v. 323, L49–L52.
4. Stauffer J. et al. *PASP*, 2010, v. 122, 885.
5. Skrutskie M. F. et al. *AJ*, 2006, v. 131, 1163.
6. Gaia Collaboration, Smart R. L. et al. Gaia Early Data Release 3: The Gaia Catalogue of Nearby Stars. arXiv:2012.02061v1 [astro-ph.SR] 3 Dec 2020.
7. Müller H. On the Cosmological Significance of Euler's Number. *Progress in Physics*, 2019, v. 15, 17–21.
8. Dombrowski K. Rational Numbers Distribution and Resonance. *Progress in Physics*, 2005, v. 1, no. 1, 65–67.
9. Müller H. The Physics of Transcendental Numbers. *Progress in Physics*, 2019, v. 15, 148–155.
10. Müller H. Global Scaling. The Fundamentals of Interscalar Cosmology. *New Heritage Publishers*, Brooklyn, New York, USA, ISBN 978-0-9981894-0-6, (2018).
11. Müller H. Fractal Scaling Models of Natural Oscillations in Chain Systems and the Mass Distribution of Particles. *Progress in Physics*, 2010, v. 6, 61–66.
12. Müller H. Scale-Invariant Models of Natural Oscillations in Chain Systems and their Cosmological Significance. *Progress in Physics*, 2017, v. 13, 187–197.
13. Khintchine A.Ya. Continued fractions. University of Chicago Press, Chicago, (1964).
14. Hilbert D. Über die Transcendenz der Zahlen e und π. *Mathematische Annalen*, 1893, v. 43, 216–219.
15. Panchelyuga V.A., Panchelyuga M. S. Resonance and Fractals on the Real Numbers Set. *Progress in Physics*, 2012, v. 8, no. 4, 48–53.
16. Müller H. Quantum Gravity Aspects of Global Scaling and the Seismic Profile of the Earth. *Progress in Physics*, 2018, vol. 14, 41–45.
17. Müller H. Global Scaling of Planetary Systems. *Progress in Physics*, 2018, v. 14, 99–105.
18. Tanabashi M. et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018), www.pdg.lbl.gov
19. Astrophysical constants. Particle Data Group, pdg.lbl.gov
20. Planck M. Über Irreversible Strahlungsvorgänge. *Sitzungsbericht der Königlich Preußischen Akademie der Wissenschaften*, 1899, v. 1, 479–480.
21. Müller H. Gravity as Attractor Effect of Stability Nodes in Chain Systems of Harmonic Quantum Oscillators. *Progress in Physics*, 2018, v. 14, 19–23.
22. Catalog of Exoplanets. Observatoire de Paris, http://exoplanet.eu/
23. Gaia EDR3, https://www.cosmos.esa.int/web/gaia/earlydr3

A Solution to the Pioneer Anomalous Annual and Diurnal Residuals

Eduardo D. Greaves[1], Carlos Bracho[2], Stephan Gift[3], and An Michel Rodriguez[4]

[1]Universidad Simón Bolívar. Apartado 89000, Caracas, Venezuela. E-mail: egreaves20002000@yahoo.com
[2]Facultad de Ingeniería, Universidad Central de Venezuela, Caracas, Venezuela. E-mail: bracho_carlos@hotmail.com
[3]Department of Electrical and Computer Engineering, Faculty of Engineering, The University of the West Indies,
St Augustine, Trinidad and Tobago, West Indies. E-mail: stephan.gift@sta.uwi.edu
[4]Universidad Simón Bolívar. Apartado 89000, Caracas, Venezuela. E-mail: anmichel.rodriguez@gmail.com

NASA's reported Pioneer 10 and 11 anomalous annual and diurnal Doppler residuals remain largely unexplained. We show they are due to the use of an invariant value of the speed of light c in the Doppler formula. The addition of the orbital speed of the Earth (~ 30 km/s) and the Earth's tangential rotational speed (~ 0.4 km/s) to the speed of light in the Doppler formula, as [18] has shown to be the velocity addition to be used, adequately fit the measured annual and diurnal Pioneer residuals. This experimentally confirms that the galilean addition of relative velocities to the speed of light satisfactorily explains the measured residuals. The newly reported values from independent analyses of the data, of the reputably constant anomalous Pioneer acceleration as a function of time, or distance from the Sun, are calculated. The values obtained, without any adjustable parameters, coincide within a percent with the experimentally measured values and are consistent with the change of the speed of light due to a decrease in the gravitational energy density with distance from the Sun as postulated by the Céspedes-Curé hypothesis. This result implies reassessment of all astronomical velocity measurements based on the Doppler Effect that have led to current cosmological theories: the Hubble constant, the expansion of the universe, the flat rotation curve of galaxies and the extreme values of the redshifts of very far away galaxies.

1 Introduction

Most of the physics related to astronomy and cosmology had been in the past based on passive astronomical observation of the measurements used to derive the theories. This is the case for Isaac Newton who derived his universal theory of gravitation from Johannes Kepler, who in turn used his own and the detailed observations of Tycho Brahe to develop his laws of planetary motion. Likewise, observations of the total Sun eclipse of 1919 by the team led by Arthur Stanley Eddington provided the first evidence in support of Einstein's General Theory of Relativity

In recent times, observational instruments have become increasingly powerful expanding visual telescopes to other ranges of the electromagnetic spectrum such as to the lower region, and to the higher regions with the radio telescopes and the x-ray and gamma ray observational satellites. These instruments have expanded our vision to ever further regions of the past history of the Universe. Moreover, with the advent of space exploration with Earth satellites and the launch of deep space probes, astronomy and cosmology now routinely utilize experimental probes to examine, refine, support or create physical theories of the cosmos. With the introduction of digital processing, computing power, extremely precise timing and the development of very high frequency electronics, accurate observations have increased to previously unforeseen ranges.

One such case is the measurement by the space agencies of extremely small phenomena that have shown minute but significant deviations from the values predicted by accepted physical theories and that have defied for lengths of time satisfactory explanation. Two examples are deviations from the predicted hyperbolic movement of space probes: the Flyby Anomaly [1, 2] and the Pioneer Anomaly [3–5]. In the Flyby Anomaly, the energy assist maneuver about the planets has been shown in several probes to deviate from the expected energy conservation prediction. In this case, speed deviations of mm/s reported are detected with errors of 10^{-2} mm/s on probes moving at speeds of several km/s.

The Pioneer Anomaly measurements of the hyperbolic movement of Pioneer 10 and 11 as well as Ulysses and Galileo have shown a minute acceleration in excess of the expected slowing towards the Sun due to its gravitational attraction. The deviations are of the order of 10^{-8}. The realization of these measurements is an extraordinary accomplishment considering that the probes are located far away in the solar system, moving at velocities in the range of several km/s. The anomalous measurements are reported with an accuracy of $\sigma_{at} = 0.32 \times 10^{-10}$ m/s^2 [5].

In addition to the assumed constant anomalous acceleration, Pioneer's Doppler residual measurements have shown annual and diurnal oscillations about the average acceleration with amplitude of about 0.8×10^{-9} m/s^2 (see Fig. 4). The magnitudes of the diurnal terms are reported to be comparable to those of the annual term. These results have been the subject of considerable discussion in the published literature: Anderson *et al* in 2002 [3, p. 40–43] concluded that they are not spacecraft-related phenomena nor artifacts of the measuring system but that they are Earth-related phenomena. In particular, the diurnal Doppler residuals exhibit a period that is close to the Earth's sidereal period.

Nieto and Anderson in 2005 [6] reported, in a very clear review, sinusoidal fits to the annual residuals showing similar values for Pioneer 10 and 11 and a phase difference of 173.2 degrees, similar to the angular separation of the two spacecrafts in ecliptic longitude.

There have been other attempts to explain the periodic anomalies. O. Olsen in 2007 [7] stated that unmodeled short-term effects are claimed to be consistent with expected values of radio plasma delay and the electron content of the Sun's Coronal Mass Ejections. Small annual and diurnal terms are considered to be artifacts of the maneuver estimation algorithm and unmodeled effects.

A. Ghosh in 2007 [8] attempted to explain these fluctuating components as due to the motions of the Earth and excess redshifts of the signal caused by velocity dependent inertial induction. He appears to be able to explain the annual and diurnal fluctuations in the anomalous acceleration of Pioneer 10.

Levi *et al* in 2009 [9] performed a data analysis independent of that of Anderson *et al* (2002), using the same data and confirming the existence of a secular anomaly. This anomaly has amplitude of about 0.8×10^{-9} m/s^2 that is compatible with that reported by Anderson *et al*. Their fit to the diurnal residuals showed the presence of significant periodic terms with the periods measured with respect to a day of 86 400 s. They reported, very accurately, periodic terms consistent with variations of one sidereal day, half a sidereal day, and half a year.

A later report on the Pioneer Anomaly by Turyshev and Toth in 2010 [5, Sec. 5.5.4, p. 86] acknowledged the presence of these oscillatory Doppler residuals ascribing them to "a mismodeling of the orbital inclination of the spacecraft to the ecliptic plane". However, in Section "7.2 Unresolved questions", it is mentioned that "Even after a best fit analysis is completed, the resulting residual is not completely random: both annual and diurnal variations are clearly visible. Is it possible to pinpoint the source of these variations?".

The current opinion (2021) that the Pioneer Anomaly was resolved as a thermal effect rests on a paper by S. G. Turyshev *et al* (2012) [10] which does a complex parametrized model for the thermal recoil force of the Pioneer spacecraft with several adjustable parameters. In particular the two adjustable parameters of Eq. (1) on page 2 predict the anomaly. However, any other parameters would negate the thermal origin of the anomaly.

Other reports that also support the thermal origin are: Rievers and Lammerzahl (2011) [12] and Francisco *et al* (2012) [13]. However, the detailed paper about the Pioneer Anomaly (55 pages of Phys. Rev. by J. D. Anderson *et al* (2002) [3]) clearly argues (see Sections VIII.B, C and D, pp. 32–35) that thermal recoil cannot account for the anomaly. Additionally, an anomaly similar to the Pioneer spacecraft was detected in Galileo spacecraft (see Section V C, p. 21) [3] and in the Ulysses spacecraft (see Section V D, p. 21) [3]. Both spacecrafts have completely different geometries and the thermal recoil theory is not applicable to them. Furthermore, the anomalous acceleration is reported to change value, decreasing rapidly and then increasing, (see discussions below, Sections 2.2 and 5.3 and references therein) during the spacecrafts' Jupiter and Saturn encounters. These reported changes of the anomaly as well as the harmonic annual and diurnal variations clearly cannot be explained by a thermal recoil theory.

More recently, L. Bilbao in 2016 [11], making use of the Vibrating Rays Theory [14], claims that relating the velocity of light and the corresponding Doppler effect with the velocity of the source at the time of detection, instead of the time of emission, it is possible to explain quantitatively and qualitatively the spacecraft anomalies. Values calculated for the annual residual approximately coincide with reported measurements for Pioneer 10 at 40 AU, $\Delta f \approx 14$ mHz and for Pioneer 11 at 69 AU, $\Delta f \approx 4.8$ mHz [11, p 310]. However, on the same arguments, the theory would predict values 5 or 6 orders of magnitude smaller than reported for the diurnal Doppler residuals measurements.

In this paper, an explanation of the constant term of the Pioneer Anomaly by Greaves in 2008 [4, 15] is reviewed with updated results and a new explanation of the oscillatory nature of the annual and diurnal Doppler residuals is presented. Both explanations are in agreement with the galilean velocity addition. The harmonic fluctuations make use of the results of analysis by Gift in 2010 of the Doppler Effect [16], in 2014 [17] and in 2017 of the Global Positioning System (GPS) [18].

2 Pioneer anomaly reported values

In order to compare the theoretical predictions with the reported values, in this section we review the literature with special emphasis on the particular phenomena pertinent to the theory presented about the anomalous acceleration values. In the light of the results below that imply different values depending on distances from the Sun and hence at the various measurement times, we do not find it surprising that a variety of values are reported.

The Pioneer Anomaly is the result of a complicated modeling procedure involving the gravitational physics predicting the probe trajectory, newtonian and relativistic, as well as a cornucopia of other phenomena such as solar radiation pressure and electromagnetic line of sight effects. The result of the modeling is compared to the measured Doppler signals received and processed by the Deep Space Network (DSN) by means of mathematical least squares fitting procedures. While there may be several possible onboard causes of anomalous results such as gas leaks or the now popular effect of asymmetric thermal radiation pressure, the different programs that have been developed generally agree on the existence of an anomaly.

We start by citing the anomalous measurements of the Galileo and Ulysses spacecrafts given by Anderson *et al* (2002) [3, p. 22, Eq. (18)]. Unfortunately, not many details are given for these anomalous measurements.

2.1 Galileo

The reported JPL values of Aerospace's analysis for the dates 2 December 1992 to 24 March 1993 give an anomalous acceleration of $a_{P(G)} = (8 \pm 3) \times 10^{-8}$ cm/s^2.

Galileo performed a second flyby of Earth at 303.1 km height at 15:09:25 UTC on 8 December 1992, adding 13,320 km/h to its cumulative speed on the way to Jupiter. Hence the reported Galileo $a_{P(G)}$ is just on or after the Earth flyby at approx. 1 AU distance from the Sun, and under full Earth and Sun gravitational acceleration field [19].

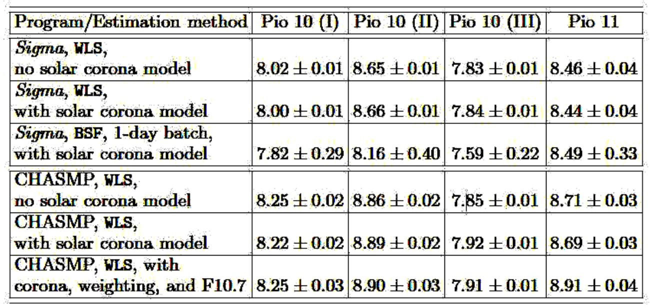
2.2 Ulysses

The JPL analysis gives $a_{P(U)} = (12 \pm 3) \times 10^{-8}$ cm/s^2. The data is from 30 March 1992 to 11 August 1994. Ulysses arrived at Jupiter on 8 February 1992 for a flyby maneuver that increased its inclination to the ecliptic by 80.2 degrees southward, away from the ecliptic plane entering an orbit around the Sun. The orbital period is approximately six years. The Sun's gravitational acceleration at the Jupiter orbit is $g_S = 2.1 \times 10^{-4}$ m/s^2, four orders of magnitude smaller than Jupiter's $g_j = 0.227$ m/s^2 gravitational acceleration at the nearest point in the flyby (4.09×10^8 m). The flyby was engineered to bring Ulysses into a Sun elliptical orbit, so that after the flyby Ulysses began movement towards the Sun with the resulting gravitational acceleration $\vec{g}_U = \vec{g}_S + \vec{g}_j$ pointing generally towards Jupiter for some time until the distance to the Sun was ≈ 5 AU. Thereafter Ulysses acceleration $\vec{g}_U$ points generally towards the Sun. Aerospace's analysis using CHASMP reports no numerical value due to measurement difficulties. However, it is stated: "The measured anomalies randomly changes sign and magnitude. The values go up to about an order of magnitude larger than a_P" [3, p 22]. This measurement and remark of Ulysses' anomalous acceleration is when the spacecraft was under Jupiter gravitational attraction just after the flyby and hence with net gravitational acceleration towards Jupiter and sometime later towards the Sun.

2.3 Pioneer 10 and 11

Table 1 of Anderson *et al* (2002) [3, p 23] reproduced below, with its original caption, gives an indication of the variability of reported values. The original data analyzed is for the following periods:

Pioneer 10: 11 years time interval (3 January 1987 to 22 July 1998), covers a heliocentric distance interval from 40 AU to 70.5 AU.



Program/Estimation method	Pio 10 (I)	Pio 10 (II)	Pio 10 (III)	Pio 11
Sigma, `WLS`, no solar corona model	8.02 ± 0.01	8.65 ± 0.01	7.83 ± 0.01	8.46 ± 0.04
Sigma, `WLS`, with solar corona model	8.00 ± 0.01	8.66 ± 0.01	7.84 ± 0.01	8.44 ± 0.04
Sigma, `BSF`, 1-day batch, with solar corona model	7.82 ± 0.29	8.16 ± 0.40	7.59 ± 0.22	8.49 ± 0.33
CHASMP, `WLS`, no solar corona model	8.25 ± 0.02	8.86 ± 0.02	7.85 ± 0.01	8.71 ± 0.03
CHASMP, `WLS`, with solar corona model	8.22 ± 0.02	8.89 ± 0.02	7.92 ± 0.01	8.69 ± 0.03
CHASMP, `WLS`, with corona, weighting, and F10.7	8.25 ± 0.03	8.90 ± 0.03	7.91 ± 0.01	8.91 ± 0.04

Table 1: Determinations of a_P in units of 10^{-8}cm/s^2 from the three time intervals of Pioneer 10 data and from Pioneer 11. As described in the text, [our Ref. 3] results from various ODP-Sigma and CHASMP calculations are listed. For ODP-Sigma, "WLS" signifies a weighted least-squares calculation, which was used with i) no solar corona model and ii) the 'Cassini' solar corona model. Also for ODP/Sigma, "BSF" signifies a batch-sequential filter calculation, which was done with iii) the 'Cassini' solar corona model. Further (see Section IX C), a 1-day batch-sequential estimation for the entire data interval of 11.5 years for Pioneer 10 yielded a result $a_P = (7.77 \pm 0.16) \times 10^{-8}$ cm/s^2. The CHASMP calculations were all WLS. These calculations were done with i) no solar corona model, ii) the 'Cassini' solar corona model, iii) the 'Cassini' solar corona model with corona data weighting and F10.7 time variation calibration. Note that the errors given are only formal calculational errors. The much larger deviations of the results from each other indicate the sizes of the systematics that are involved. (Acronyms are: ODP - JPL's Orbit Determination Program; CHASMP - Aerospace Corporation's Compact High Accuracy Satellite Motion Program.)

Pioneer 11: data of 3 years (5 January 1987 to 1 October 1990), covers a heliocentric distance interval much closer to the Sun, from 22.42 to 31.7 AU.

Additionally, Anderson *et al* (2002) [3, on p. 27] quote: "For Pioneer 10, two different analysis programs: Sigma and CHASMP (*measurements*) are similar, 7.82×10^{-8} cm/s^2 and 7.89×10^{-8} cm/s^2, the weighted average of these two to yield $a_{Pio10} = (7.84 \pm 0.01) \times 10^{-8}$ cm/s^2 (experimental).

"For Pioneer 11, we only have the one 3 3/4 year data arc. The weighted average of the two programs' no corona results is $(8.62 \pm 0.02) \times 10^{-8}$ cm/s^2."

2.3.1 Information of planetary encounters

The Pioneer 10 original data spans heliocentric distance interval from 40 AU to 70.5 AU, as mentioned above. Hence it does not include the Jupiter flyby at 5.2027 AU on 1974.

Pioneer 11's original data covers a heliocentric distance interval from 5.80 to 29.50 AU. It includes just after the Jupiter flyby at 5.2027 AU and the Saturn encounter at 9.546 AU on 1979. Also near encounter with Uranus at 19.2 AU on approx. 1986 and with Neptune at 30.09 AU on approx. 1990. Moreover, a report in 2005 of Nieto and Anderson [6] pro-

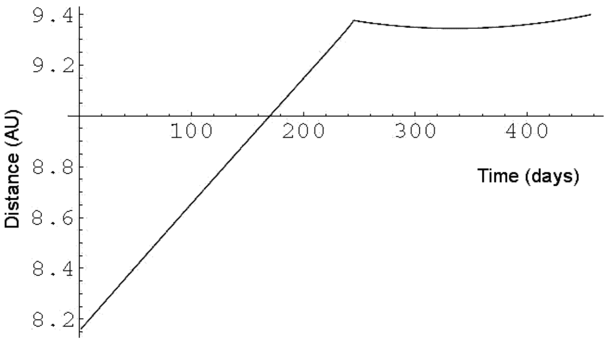


Fig. 1: "A plot of Pioneer 11's distance from the Sun (in AU) vs time (in days starting with 1 Jan. 1979) near Saturn encounter (on Sept. 1, 1979)" [6, p 14]. *Captions of axes added.*

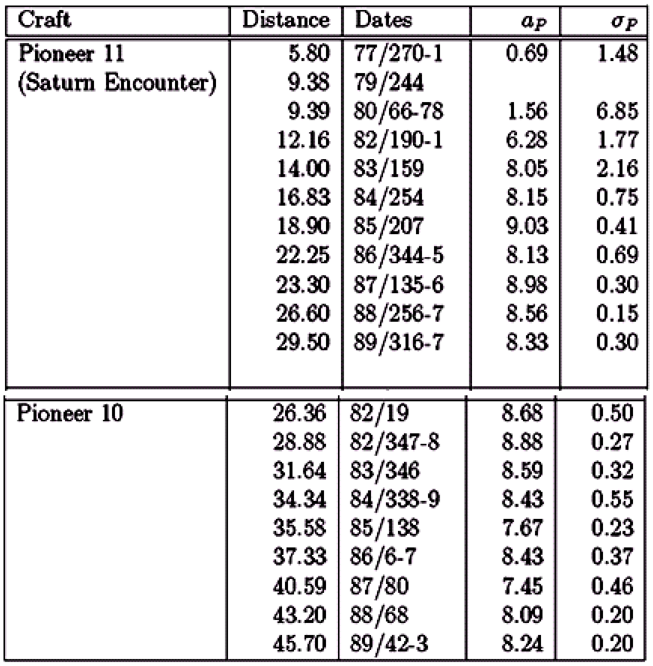
vides additional insight to the planetary encounters and the harmonic residuals. They report that the initial two points in the Pioneer 11 anomalous acceleration shown in their Figures 4 and 5 (our Fig. 2) were near the distances of Jupiter and Saturn encounters. They provide a figure showing the distance (AU) as a function of time (in days) around the Saturn flyby of Pioneer 11. See Fig. 1 with its original caption.

We find this figure very illuminating as at these times the spacecraft was under the gravitational attraction of Saturn and perhaps also under the influence of its higher space energy density as discussed below.

Regarding the annual residuals, Nieto *et al* [6] mention on p. 14:

> Plots of the anomaly versus time were also made from these data points. These showed, as might be suspected from Figures 4 and 5, [our Fig. 2] a possible annual variation. This observation would be a forerunner of the discussion in Section IX-C of [12], [our Ref. [3]]. Doing fits to the data points, the best estimate of the amplitude of the Pioneer 10 sine wave is $(0.525 \pm 0.155) \times 10^{-8}$ cm/s^2 and that of the Pioneer 11 wave is $(0.498 \pm 0.176) \times 10^{-8}$ cm/s^2 (here with the first three points omitted). The sine waves seem real, with, e.g. a 95 percent probability that the Pioneer 10 amplitude lies between 0.199 and 0.834×10^{-8} cm/s^2. The difference in phase between the Pioneer 10 and Pioneer 11 waves is 173.2 degrees, similar to the angular separation of the two spacecrafts in ecliptic longitude. [This is 204.28 degrees at the present time.]

Useful information is provided in Table II which contains the numerical data for each spacecraft containing the distance, dates, reported anomalous acceleration a_P and the error ΔP. Using this information, we find it helpful to plot the reported dates and distances (see Fig. 3) as this information allows the determination of the distance or dates of reported a_P when the information is not given.



Craft	Distance	Dates	a_P	σ_P
Pioneer 11 (Saturn Encounter)	5.80	77/270-1	0.69	1.48
	9.38	79/244		
	9.39	80/66-78	1.56	6.85
	12.16	82/190-1	6.28	1.77
	14.00	83/159	8.05	2.16
	16.83	84/254	8.15	0.75
	18.90	85/207	9.03	0.41
	22.25	86/344-5	8.13	0.69
	23.30	87/135-6	8.98	0.30
	26.60	88/256-7	8.56	0.15
	29.50	89/316-7	8.33	0.30
Pioneer 10	26.36	82/19	8.68	0.50
	28.88	82/347-8	8.88	0.27
	31.64	83/346	8.59	0.32
	34.34	84/338-9	8.43	0.55
	35.58	85/138	7.67	0.23
	37.33	86/6-7	8.43	0.37
	40.59	87/80	7.45	0.46
	43.20	88/68	8.09	0.20
	45.70	89/42-3	8.24	0.20

Table 2: Pioneer 11 and 10 early data points (Distance in AU, Date year/days-of-year, anomaly a_P and error σ_P in units of 10^{-8} cm/s^2 from [6].

Toth and Turyshev in 2007 [20, p. 15] comment results found during the Jupiter–Saturn cruise phase: "Right at the time of the Saturn encounter, however, when the spacecraft passed into an hyperbolic escape orbit, *there was a rapid increase in the anomaly, whereafter it settled into the canonical value*" [our emphasis].

2.3.2 Independent analysis of Pioneer data

There have been several further independent analyses of the original data which were made available since 2002 and are reviewed below.

C. Markwardt (2002) [21] performed an independent analysis of radio Doppler tracking data from the Pioneer 10 spacecraft for the time period 1987–1994. His best-fit value for the acceleration, including corrections for systematic biases and uncertainties, is $(8.60 \pm 1.34) \times 10^{-8}$ cm/s^2, directed towards the Sun.

O. Olsen (2007) [7] does an independent analysis of the Pioneer 10 and 11 data using the HELIOSAT program developed by one of the authors at the University of Oslo. The data used spans the three periods defined by Anderson *et al* (2002) for Pioneer 10: Interval I spans 1 January 1987 to 17 July 17 1990, Interval II spans 17 July 1990 to 12 July 12 1992 and Interval III continues up to 21 July 1998. The anomalous accelerations ($\times 10^{-8}$ cm/s^2) obtained are given in his Table I from which we extract: Pioneer 10: Interval I = 7.85 ± 0.02;

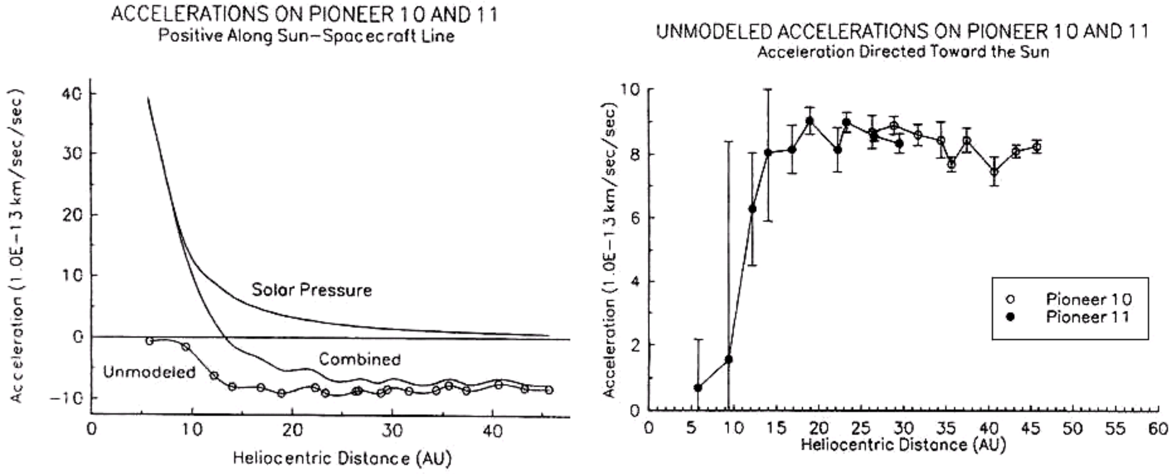


Fig. 2: Left: Accelerations on the Pioneer spacecraft. Right: Anomalous acceleration as a function of distance [3, p. 19].

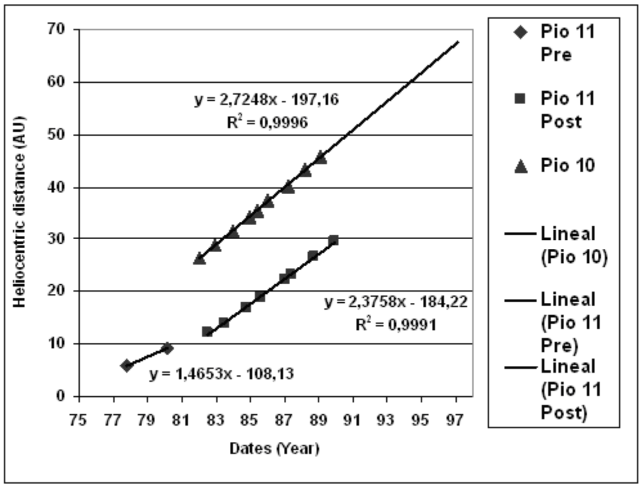


Fig. 3: Heliocentric distance (AU) as a function of dates (year) for Pioneer 10 and 11 positions. Lines are linear fits. For Pioneer 11 pre and post Jupiter flyby. (Data from Table II.)

Interval II = 8.78 ± 0.01; Interval III = 7.75 ± 0.01; Pioneer 11 = 8.10 ± 0.01. From the paper's conclusions: "The unmodeled short term effects are claimed to be consistent with expected values of radio plasma delay and the electron content of Coronal Mass Ejections. Small annual and diurnal terms are ascribed as artifacts of the maneuver estimation algorithm and unmodeled effects acting on the spacecraft or on the radio transmissions."

V. T. Toth does an independent analysis of the orbit of the Pioneer spacecrafts reporting in 2009 [22, p. 18] for Pioneer 10* $a_{P10} = (9.03 \pm 0.86) \times 10^{-8}$ cm/s^2 (period 1987 to 1998) and for Pioneer 11 $a_{P11} = (8.21 \pm 1.07) \times 10^{-8}$ cm/s^2 (period 1986 to August 1990). Toth also attempted in his analysis to test the extent to which the anomalous acceleration is constant in time. To this end, he implemented the ability to estimate a secondary acceleration, i.e. "jerk" term in the orbital solution.

The results obtained were: for Pioneer 10, $a_{P10} = (10.96 \pm 0.89) \times 10^{-8}$ cm/s^2 [3, p. 20], with a variation rate of $da_{P10}/dt = -(0.21 \pm 0.04) \times 10^{-6}$ cm/s^2/year and for Pioneer 11, the result was $a_{P11} = (9.40 \pm 1.12) \times 10^{-8}$ cm/s^2, with a variation rate of $da_{P11}/dt = -(0.34 \pm 0.12) \times 10^{-8}$ cm/s^2/year. Toth goes on to state: "an anomalous acceleration that is a slowly changing function of time (*decreasing*) cannot be excluded at present" [our italics].

Levi *et al* in 2009 [23] performed a data analysis independent from that of Anderson *et al* (2002) using the same Pioneer 10 data confirming the existence of a secular anomaly with an amplitude of about 8×10^{-8} cm/s^2 compatible with that reported by Anderson *et al* (2002) and providing additional insight into the phenomenon.

2.4 Annual and diurnal Doppler residuals

The first indication of the oscillatory nature of the Pioneer Anomaly came from an examination of the data in Fig. 2. The observations are addressed in detail in Anderson *et al* (2002) [3, pp. 40-41]. From that report, we show Figs. 4, 5 and 6.

Levi *et al* in 2009 [9], performed an important and illuminating independent analysis of the diurnal periodic terms during a short time span of (we quote): "23 November 1996 to

---

[*] Toth and Levi *et al* express all values in SI units. We have converted accelerations to cm/s^2 as used in most Pioneer reports.

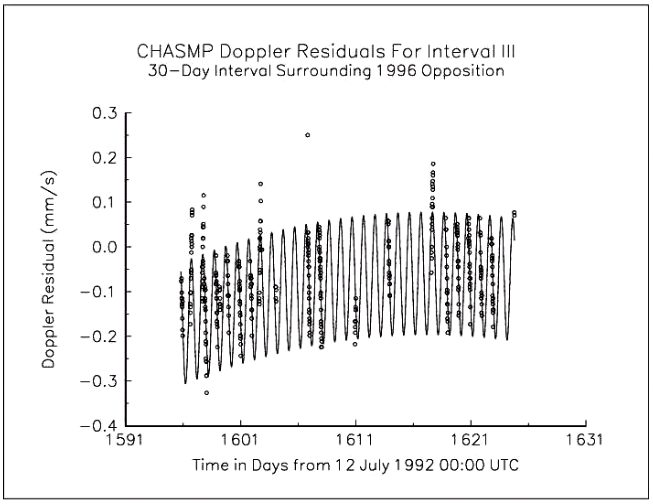


Fig. 4: Diurnal residuals. "CHASMP acceleration residuals from 23 November 1996 to 23 December 1996" [3, Fig. 18, p. 41].

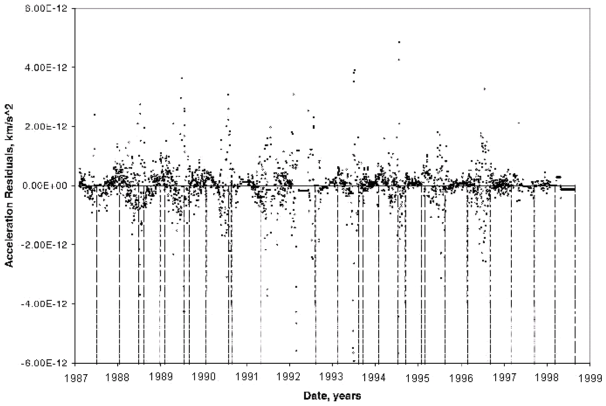


Fig. 5: Annual residuals. "ODP 1-day batch-sequential acceleration residuals using the entire Pioneer 10 data set. Maneuver times are indicated by the vertical dashed lines". [3, Fig. 17, p. 40].

23 December 1996 when Pioneer 10 was on opposition (Sun, Earth and Pioneer 10 aligned in this order). This data set is thus less affected by solar plasma and it shows daily variations of the residuals". The analysis of residuals shows the presence of significant diurnal periodic terms with the periods measured with respect to a day = 86 400 s. Their spectral analysis of the periodic terms yields the following measured periods: T1 = 0.9974 ± 0.0004 day, T2 = (1/2)(0.9972 ± 0.0004) day and T3 = 189 ± 32 days. "As T = 0.997 day = 1.0 sidereal day, these periods are consistent, (*within 0.02 percent*), with variations of one sidereal day, half a sidereal day, and half a year." (Year/2 = 182.5 days) [Our italics]. These results clearly indicate that the periodic terms in the Doppler residuals are not produced by on-board phenomena or due to solar corona affecting transmission signals, but rather that they are intimately related to Earth movement relative to the Pioneer position in the sky. To illustrate their results we reproduce below (Fig. 7) a section of Figure 3 in that report.

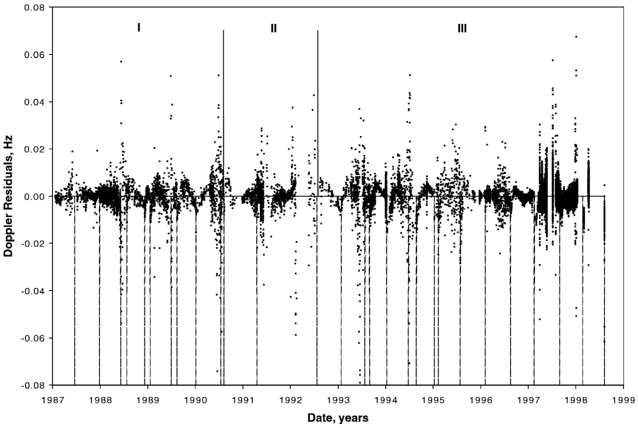


Fig. 6: "ODP Doppler residuals in Hz for the entire Pioneer 10 data span. The two solid vertical lines in the upper part of the plot indicate the boundaries between data Intervals I/II and II/III, respectively. Maneuver times are indicated by the vertical dashed lines in the lower part of the plot." [3, Fig. 13, p. 25].

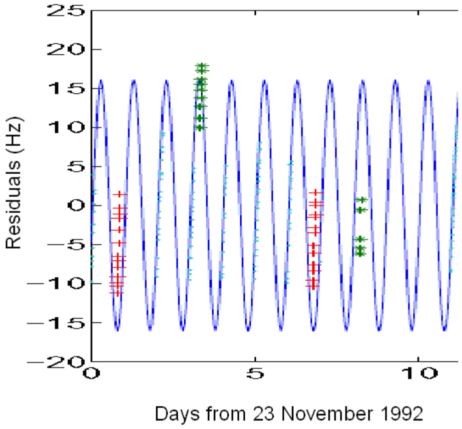


Fig. 7: Fitted residuals of the Doppler tracking data of Pioneer 10, for a 10-day period near opposition. Different symbols or colors refer to different couples of receiving stations [6, expanded section of Figure 3, p. 6]

All of the reports shown above use the original data and do not include the early stage of the Pioneer missions. It has been stated in several reports [6, 20, 24] the convenience to recover and analyze the data from the beginning of the missions. A very commendable effort has been made to recover the earlier data, which after considerable effort, has been secured in modern digital media. A detailed report of this contribution is found in [22, p. 4]. However, to our knowledge, the required detailed analysis of the earlier stages has not been reported.

At the present times, Pioneer 10 is in the constellation of Taurus. The current Right Ascension of Pioneer 10 is 05h 16m 17s and the Declination is +26°02′40″. Pioneer 11 is currently in the constellation of Scutum. The current Right Ascension of Pioneer 11 is 18h 53m 32s and the Declination is −08°42′43″ [25].

3 Pioneer anomaly "constant" term

In this section, we review the theory of the calculation of the Pioneer 10 and 11 anomalous "constant" term. We start with the statement of the Céspedes-Curé Hypothesis [26, p. 279], [4, 27–29] Eq. (1): The speed of light on Earth's surface at 1 AU (S Sun, E Earth) is given by

$$c = \frac{k}{\sqrt{\rho}} = \frac{k}{\sqrt{\rho^* + \rho_S + \rho_E}} \tag{1}$$

where k is a proportionality constant and ρ the energy density in J/m^3 on the surface of the Earth which is a sum of the contribution of the constant energy density due to far away stars and galaxies ρ^* and the constant values due to the Sun ρ_S and Earth ρ_E given by (2) below. Calculation shows that the contribution of the Moon and other planets are negligible.

The energy density of a mass [26, p. 163], [2, Eq. (2), p. 50], [4] is given by

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

where G is Newton's gravitational constant, M is the mass and r is the distance from the mass center. Eq. (2) shows the energy density of a mass decreases very rapidly due to the r exponent of 4 in the denominator.

The speed of light far away from Earth and the Sun, at Pioneer's position, is given by

$$c' = \frac{k}{\sqrt{\rho'_{far}}}\,. \tag{3}$$

Here ρ'_{far} is the energy density at the site of Pioneer. In (3), ρ'_{far} contains a sum of the gravitational energy density of the far away stars and galaxies ρ^*, the Sun's and the energy density of other planets, which are relatively near in the spacecraft's trajectory towards outer space. These include the Earth in the very early stage of the mission and any planets during flyby or relative close approach, which includes the Jupiter flyby, the Saturn flyby and possibly near encounters to other planets. Hence

$$\rho'_{far} = \rho^* + \frac{G}{8\pi}\sum_{i=1}^{n}\frac{M_i^2}{r_i^4}\,. \tag{4}$$

Figs. 8 and 9 shown below give an indication of these encounters. A close look at these figures clearly shows that the gravitational energy density and gravitational acceleration along the trajectory of Pioneer 10 and 11 are different predicting different values of the anomalous acceleration as is reported.

The index of refraction of space, relative to the vacuum index on Earth, at Pioneer's position is obtained using (1) and (3):

$$n' = \frac{c}{c'} = \frac{\sqrt{\rho'_{far}}}{\sqrt{\rho}}\,, \tag{5}$$

so that the speed of light far away is:

$$c' = c\,\frac{\sqrt{\rho}}{\sqrt{\rho'_{far}}}\,. \tag{6}$$

Eq. (6) implies that c' is *greater* than c and increases with distance as ρ'_{far} decreases with distance. However, c' becomes almost constant when Pioneer goes past the planets and their energy density becomes negligible. The Sun's contribution continues to decrease leaving ρ^*, the constant energy density of far away stars and galaxies

Spacecraft velocity and accelerations are measured basically with the Doppler formula $\Delta f = f\,(v/c)$ where f would be a spacecraft-generated signal. However, Pioneer spacecraft did not have an accurate oscillator onboard. Communication uplink from Earth is at ∼ 2.11GHz. The spacecraft's coherently received signal is accurately multiplied by the (240/221) ratio and signals beamed at approximate downlink frequency 2.295 GHz. The signals are sent and received by the Deep Space Network (DSN) and processed in the manner described in detail by Anderson *et al* [3, pp. 7–12]. In this manner, the observable is a very precise Doppler shifted frequency $\Delta f = (f/c)(dl/dt)$ [3, p. 9, Eq. (1)], where l is the overall optical distance. In our notation $v = (dl/dt)$ so that the spacecraft speed is obtained with:

$$\vec{v} = \frac{c\Delta f}{f}\,\frac{\vec{r}}{|r|}\,. \tag{7a}$$

Differentiating (7a) with respect to time, the measured spacecraft acceleration is

$$\vec{a}_{JPL} = \frac{d\Delta f}{dt}\,\frac{c}{f}\,\frac{\vec{r}}{|r|}\,. \tag{7b}$$

Here Δf is the shift of the frequency f and $\frac{d\Delta f}{dt}$ the measured drift of the frequency due to the Pioneer acceleration produced by gravitation at the spacecraft site, mainly due to the Sun. $\vec{a}_{JPL}$ is a derived acceleration vector in the direction of the gravitational force causing it. Examination of (7a) and (7b) shows that, if the velocity of light c is not invariant but rather given by (6) as proposed in this work, measurement of velocity and acceleration in locations of space with lower

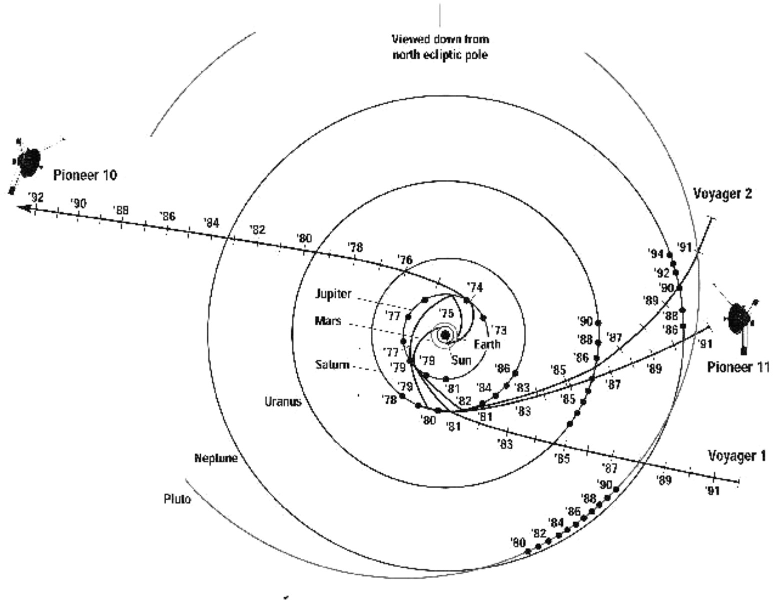


Fig. 8: "Ecliptic pole view of Pioneer 10, Pioneer 11, and Voyager trajectories. Pioneer 11 is traveling approximately in the direction of the Sun's orbital motion about the galactic center. The galactic center is approximately in the direction of the top of the figure." [3, p. 5].

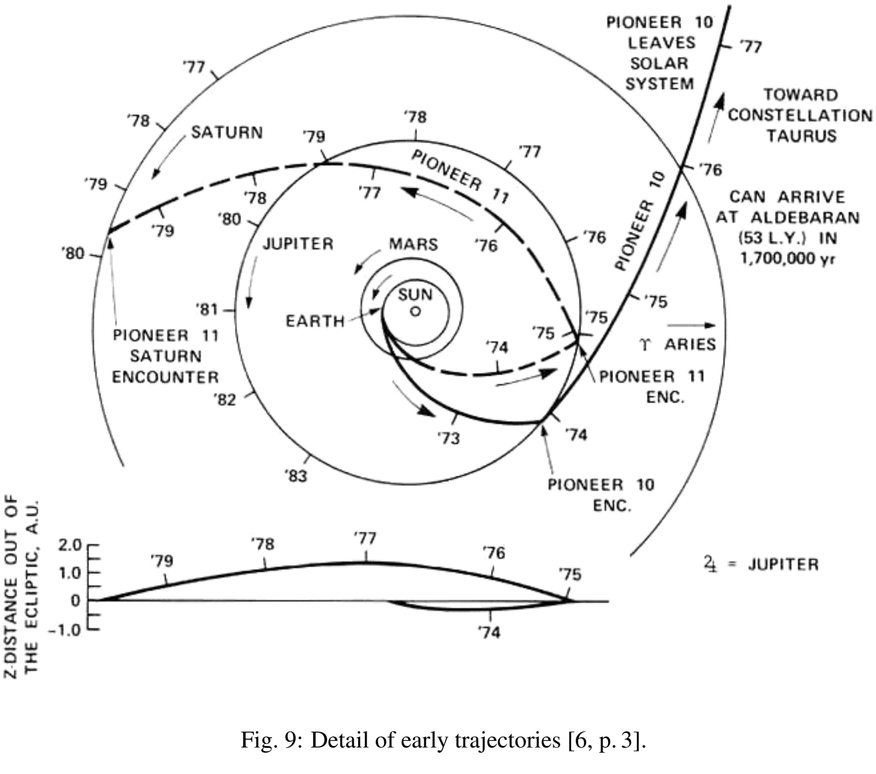


Fig. 9: Detail of early trajectories [6, p. 3].

gravitational energy density than on Earth's surface, both result in overestimation of these quantities. This leads to the belief that an anomalous acceleration towards the sun is acting.

At this point it is instructive to mention that c' differs very little from c and the magnitudes of n', the index of refraction of space, that are predicted with (5) are very nearly equal to 1. The values of n' on the surface of planets differ from Earth's

Planet	n'	Planet	n'
Mercury	0.99997382	Saturn	100.000.349
Venus	0.99999527	Uranus	0.99999524
Earth	100.000.000	Neptune	100.000.737
Mars	0.99997758	Pluto	0.99997385
Jupiter	100.014.145	Moon	0.99997454

Table 3: Values of the vacuum index of refraction n' on the surface of the planets and the Moon. The value of $\rho^* = 1.09429 \times 10^{15}$ J/m^3 calculated by Céspedes-Curé [26, p. 279] was used in evaluating n' with (5).

by very little. Table III from [4] shows the results of calculating n' with the use of (5). The values of the planets are close to 1.0 being caused by the local gravitational energy density being not too different from the surface of the Earth.

The correct value of Pioneer's acceleration is obtained with Newton's gravitational force:

$$\vec{a}_N = G \sum_{i=1}^{n} \frac{M_i}{r_i^2} \frac{\vec{r}}{|r|} . \tag{8}$$

Here the acceleration of gravity (i = Sun and planets) is mainly from the Sun, but in the early stages of the mission it will be affected by other planets which are relatively near during energy assist maneuvers (flyby) or near encounter in its trajectory towards outer space.

The Pioneer acceleration is measured with the Doppler formula (7b) with the accepted value c of the speed of light and the uplink $f = 2.113$ GHz frequency. If instead of c we use the speed of light c' given by (6), we get a corrected Doppler-derived acceleration:

$$\vec{a}' = \frac{d\Delta f}{dt} \frac{c}{f} \frac{\sqrt{\rho}}{\sqrt{\rho'_{far}}} \frac{\vec{r}}{|r|} . \tag{9}$$

The difference between the acceleration $\vec{a}'$ as proposed here in (9) and $\vec{a}_{JPL}$ calculated with (7b) gives the predicted anomalous acceleration:

$$\vec{a}_P = \frac{d\Delta f}{dt} \frac{c}{f} \left(\frac{\sqrt{\rho}}{\sqrt{\rho'_{far}}} - 1 \right) \frac{\vec{r}}{|r|} . \tag{10}$$

4 Pioneer annual and diurnal residuals

Here, we present the theory to calculate the harmonic Doppler residuals of the Pioneer 10 and 11 spacecraft. Due to Earth's rotation and translation, the measured acceleration a_{JPL}, has superimposed Doppler shifts due to the components of these movements in the direction of the spacecraft. They are incorporated in the models used to predict the spacecraft movement by the standard galilean addition of velocities, to the predictions of gravitational theory.

We treat first the effect of Earth's rotation. Let V_{ER} be the equatorial tangential velocity (≃0.4 km/s). At the latitude λ of the DSN antennas, the tangential velocity in the direction of Pioneer changes by $\cos\lambda$. As the day progresses, the component decreases by the factor $\cos(\omega_R t + \phi_R)$, where ω_R is the Earth's *sidereal* angular rotation velocity and ϕ_R an Earth rotational phase angle. Hence the rotational Earth's velocity in the direction of Pioneer is

$$v_E = v_{ER} \cos\lambda \cos(\omega_R t + \phi_R) . \tag{11}$$

For argument's sake, we take $(\omega_R t + \phi_R)$ to be equal to 0 degrees when Pioneer is just in the East of the DSN station. Then $\cos(\omega_R t + \phi_R) = 1$ and the velocity predicted is maximum when Pioneer is in the East horizon of the DSN antenna. The component reaches a null value when Pioneer is in the zenith of the DSN station $(\omega_R t + \phi_R) = 90^{\circ}$, and becomes negative, reaching a maximum negative value when it is exactly in the West sky of the DSN station. This component has to be added to the speed of light in (10).

In regards to Earth's translation about the Sun, let v_{ET} be Earth's translation velocity (approx. 30 km/s). The component of the translation velocity in the direction of Pioneer is

$$v_E = v_{ET} \cos\lambda \cos(\omega_T t + \phi_T) \tag{12}$$

with ω_T the Earth's sidereal angular translational velocity about the Sun and ϕ_T an Earth translational phase angle.

This component is a maximum when Pioneer is in quadrature and becomes null when it is in opposition (Sun, Earth, Pioneer alignment) or in conjunction with the Sun (Earth, Sun, Pioneer alignment). See Fig. 10.

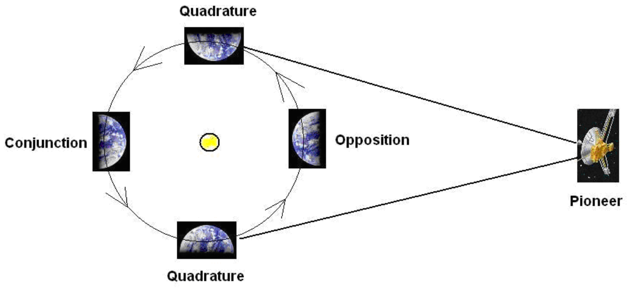


Fig. 10: The Earth translation under the position of Pioneer.

To calculate the annual and diurnal Doppler residuals, we use the galilean velocity addition as demonstrated by Gift [18] adding *to the speed of light* in (10) the Earth's orbital translation and rotational velocity

$$\begin{aligned} \vec{a}_P = \frac{d\Delta f}{dt} \frac{1}{f} \Big(& c + v_{ET} \cos\lambda \cos(\omega_T t + \phi_T) + \\ & + v_{ER} \cos\lambda \cos(\omega_R t + \phi_R) \Big) \left(\frac{\sqrt{\rho}}{\sqrt{\rho'_{far}}} - 1 \right) \frac{\vec{r}}{|r|} . \end{aligned} \tag{13}$$

In order to calculate $\vec{a}'_P$ with (13) as a function of distance, it is necessary to know the frequency drift $d\Delta f/dt$. The value of the frequency drift for different distances is not available. It has been measured and considered to be constant for the later part of the trajectories. Nevertheless, its value is important for calculating the earlier part of the trajectories. We see that it is possible to calculate values by equating the correct newtonian acceleration given by (8) to the measured acceleration given by (7b). Solving for the frequency drift we obtain:

$$\frac{d\Delta f}{dt} = \frac{f}{c}\, G \sum_{i=1}^{n} \frac{M_i}{r_i^2}\,. \tag{14}$$

Here we have to use an invariant c as used by JPL.

On page 16 of Anderson *et al* (2002), it is stated that the measured frequency in Hz is converted to Doppler velocity by the use of their Eq. (13), namely $v = c\,\Delta f/2f$ in our notation. This indicates that the values reported are obtained using a *double* Doppler (uplink–downlink) velocity. Hence our formulation for the frequency shift has to be multiplied by a factor of 2:

$$\frac{d\Delta f}{dt} = \frac{2f}{c} G \sum_{i=1}^{n} \frac{M_i}{r_i^2}\,. \tag{15}$$

With (15) in (13) we get:

$$\begin{aligned} \vec{a}'_P = 2\bigg[1 &+ \frac{v_{ET}}{c}\, \cos\lambda \cos(\omega_T\, t + \phi_T) + \\ &+ \frac{v_{ER}}{c}\, \cos\lambda \cos(\omega_R\, t + \phi_R)\bigg] + \\ &+ G \sum_{i=1}^{n} \frac{M_i}{r_i^2} \left(\frac{\sqrt{\rho}}{\sqrt{\rho'_{far}}} - 1 \right). \end{aligned} \tag{16}$$

Eq. (16) predicts both the constant term of the Pioneer anomalous acceleration towards the Sun and the smaller harmonic Doppler residuals in units of acceleration (m/s^2). It predicts different values for Pioneer 10 and 11 with the differences notably contained in the gravitational acceleration acting on the spacecraft (particularly during planetary encounters in the early phase of the missions). This difference is also due to the distances contained in the ρ'_{far} factor, and the different phase angles ϕ_T for the two spacecrafts.

Since they are going in different directions in the ecliptic plane, the difference $\Delta\phi = \phi_{TPio10}$ - ϕ_{TPio11} should be equal to the difference of their Right Ascensions. This is a variable quantity during the early phase of the mission. However, at the present time, it is nearly constant and equal to (Pioneer 10: 05h 16m 17s) – (Pioneer 11: 18h 53m 32s) = 204.3 degrees.

The much higher translational velocity v_{ET} of Earth with an annual period dominates over the smaller diurnal variations of $\vec{a}'_P$. However the annual variations are slow and the Earth–Pioneer component of v_{ET} is very small during conjunction and opposition.

Figure 4 from Anderson *et al* (2002) [3] clearly shows the harmonic Doppler residuals after subtracting the constant term. These figures are made up of very many different measurements without any established periodicity. Measurements were made when the probe was in the sky of one of the DSN station antennas at arbitrary times of the rotational and translational positions of Earth, which means for (16), different values of the rotational and translational phase angles ϕ_R and ϕ_T.

There are 3 DSN Stations located approximately 120 degrees apart (Madrid, Spain, Goldstone, California and Camberra, Australia). This means that measurements from each station differ in phase angle ϕ_R by about 120 degrees so that in general, it is not expected that Doppler residuals exhibit an oscillatory continuity for any length of time. Nevertheless, as mentioned and reviewed in Section 2 above, previous workers have made detailed analyses of these harmonic Doppler residuals taking into account the phase differences.

We may also derive the Pioneer annual and diurnal Doppler residuals in units of velocity or alternatively in units of frequency as has been reported [3, 9, 22].

The Doppler formula is

$$\Delta f = \frac{v_P}{c}\, f \tag{17}$$

with v_p the speed of the Pioneer spacecraft, f the transmitting frequency, Δf the frequency change and c the speed of light considered a constant. In the case of the Pioneer spacecraft, it is a "Double" Doppler effect as mentioned above, hence:

$$\Delta f = 2\, \frac{v_P}{c}\, f\,. \tag{18}$$

If, instead of c, we use c plus the Earth speed following the results of Gift (2017) [18], then we write

$$\Delta f'' = 2\, \frac{v_P}{c + v_E}\, f\,. \tag{19}$$

NASA expects (18) and gets Δf plus or minus a "residual" which we think is due to not using (19). Hence the residual must be (18) minus (19):

$$\Delta f'' = \Delta f - \Delta f' = 2 v_P f \left(\frac{1}{c} - \frac{1}{c + v_E} \right). \tag{20}$$

Or

$$\Delta f'' = 2 v_P f \left(\frac{v_E}{(c^2 + c v_E)} \right).$$

This approximates to

$$\Delta f'' \approx 2 v_P f \left(\frac{v_E}{c^2} \right). \tag{21}$$

This relation gives the maximum values. To calculate the diurnal Doppler residuals as a function of time, we substitute (11) in (21):

$$\Delta f''_D \approx 2f\, \frac{v_P\, v_{ER}}{c^2} \cos\lambda \cos(\omega_R\, t + \phi_R)\,. \tag{22a}$$

The equivalent relation for annual residuals is obtained by substituting (12) in (21)

$$\Delta f''_A \approx 2f\,\frac{v_P v_{ET}}{c^2}\cos\lambda\cos(\omega_T\,t+\phi_T)\,. \tag{22b}$$

The result (22) gives the annual and diurnal residuals $\Delta f''$ in units of frequency (Hz). We want to compare with results in velocity units such as (mm/s) as shown in Fig. 4. To convert from Hz to m/s Anderson *et al* (2002) [3, p. 16] uses

$$\Delta v'' = \frac{\Delta f'' c}{2f}\,. \tag{23}$$

Then substituting (22) in (23) we get for the diurnal Doppler residuals in [m/s]:

$$\Delta v''_D = \frac{v_P\,v_{ER}}{c}\cos\lambda\cos(\omega_R\,t+\phi_R)\,. \tag{24a}$$

The equivalent relation for annual residuals is

$$\Delta v''_D = \frac{v_P\,v_{ET}}{c}\cos\lambda\cos(\omega_T\,t+\phi_T)\,. \tag{24b}$$

5 Results

In this section we use the theory developed above to predict qualitatively and quantitatively the reported Pioneer Anomaly "constant" and harmonic Doppler residuals.

5.1 Pioneer 10 anomaly at 20 AU

At 20 AU we calculate the anomalous acceleration with (16). For this "constant" term, we omit the terms dealing with the harmonic Doppler residuals and consider only the gravitational acceleration and energy density (in ρ'_{far}) due to the Sun and Earth:

$$\vec{a}_P = 2G\left(\frac{M_S}{r_S^2}+\frac{M_E}{r_E^2}\right)\left(\frac{\sqrt{\rho}}{\sqrt{\rho'_{far}}}-1\right)\frac{\vec{r}}{|r|}\,. \tag{25}$$

This expression predicts:

$$a_P = 7.754\times10^{-8}\ \mathrm{cm/s^2}\,. \tag{26}$$

This calculated value differs by just 1.2 percent from the value $a_P = 7.85 \pm 0.02\times10^{-8}\ \mathrm{cm/s^2}$ reported by O. Olsen (2007) [7] in an independent analysis of the Pioneer 10 data for Interval I. The value calculated in (26) also coincides, within the error estimation, with the result quoted by Anderson *et al* (2002) [3, p. 24]: "1-day batch-sequential estimation for the entire data interval of 11.5 years for Pioneer 10 (*which*) yielded a result $a_P = (7.77\pm0.16)\times10^{-8}\ \mathrm{cm/s^2}$." In this case our calculation differs by just −0.2 percent.

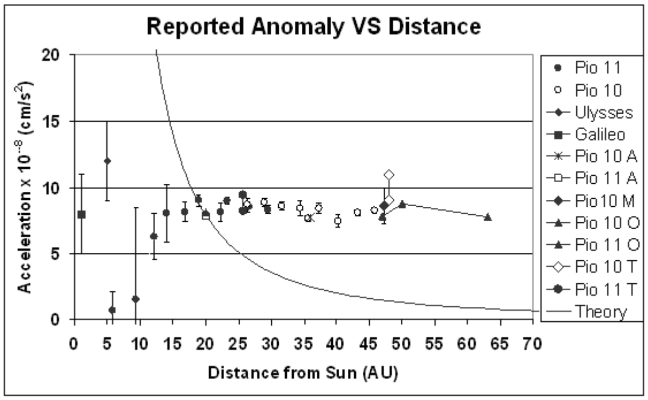


Fig. 11: Anomalous acceleration × 10^{-8} in units of (cm/s^2) as a function of distance from the Sun. Values of anomalous acceleration reported and reviewed above are plotted with the theoretical line according to (25) (A Anderson, M Markwardt, O Olsen, T Toth).

5.2 Pioneer anomaly as a function of distance from the Sun

To present the anomalous acceleration predicted as a function of distance, we show below results of a simple model with the influence of the Sun and Earth ignoring the other planets.

The theoretical curve in Fig. 11 shows a variable slope decreasing with distance. V. Toth (2009) reports in his independent analysis, as quoted above, values for a_P variation rates for Pioneer 10 and 11. However, it is not stated for what distances or dates are these quantities deduced. The value reported for Pioneer 10 [22, p 20] is $da_{P10}/dt = -(0.21\pm0.04)\times10^{-6}\ \mathrm{cm/s^2/year}$. We find that the theoretical curve in Fig. 11 exhibits that slope exactly, within the uncertainty shown, at a distance between 42 and 48 AU.

For Pioneer 11, the Toth reported variation rate is $da_{P11}/dt = -(0.34\pm0.12)\times10^{-8}\ \mathrm{cm/s^2/year}$. We find that the theoretical curve in Fig. 11 exhibits that slope exactly, within the uncertainty shown, at distances between 29 and 38 AU. Hence, we fully agree with Toth's comment: "an anomalous acceleration that is a slowly changing function of time (*decreasing*) cannot be excluded at present" [our italics].

5.3 Pioneer anomaly during Jupiter flyby

Ulysses, Pioneer 10 and 11 had close encounters with Jupiter as part of mission exploration objectives and for orbit modifications or energy assistance maneuvers. We show now how the theory developed here can explain some of the observations reported during Jupiter flyby by these spacecrafts. The effects of the gravitational energy density due to the planets are very short range according to (2) and even for the Sun [4]. Likewise the gravitational acceleration produced by the planets is relatively short range compared to the large distances traversed by these spacecrafts. To put the values in perspective, we show in Fig. 12 the gravitational acceleration of the

Sun and the planets each centered about their orbital distance to the Sun.

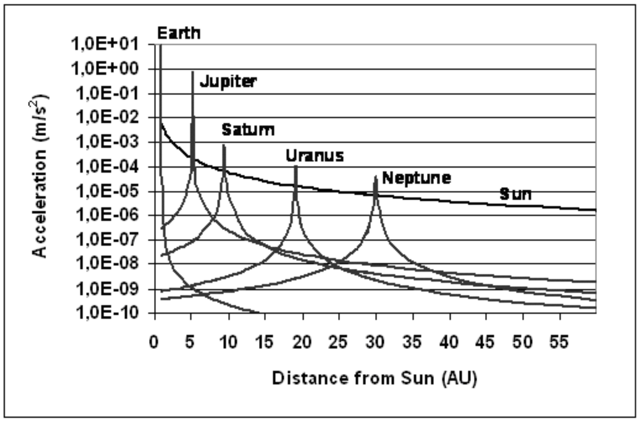


Fig. 12: Gravitational acceleration of the Sun, Earth and planets centered about the position of their orbits about the Sun. (Values are calculated to within 0.2 AU of each planet, and centered at nominal orbital distance).

Fig. 12, top line, shows that compared to the gravitational acceleration of the Sun the planet's acceleration affects only their immediate vicinity. If we rewrite (16) considering just the Sun and Jupiter and emphasizing the vectorial character of $\vec{a}_P$, we get

$$\vec{a}_P = 2G\left(\frac{\sqrt{\rho}}{\sqrt{\rho'_{far}}} - 1\right)\left(\frac{M_S}{r_S^2}\frac{\vec{r}}{|r_S|} + \frac{M_J}{r_J^2}\frac{\vec{r}}{|r_J|}\right). \qquad (27)$$

Figs. 13 and 14 show the Jupiter flybys of the Pioneer spacecraft. Judging from the incoming and outgoing trajectories towards Jupiter in the polar view of the Pioneer 11 flyby, we deduce that the resulting vectorial gravitational acceleration due to Jupiter and the Sun was mainly in the direction of the Sun, but with the gravitational attraction of Jupiter in the opposite direction. Hence the initial two points in the Pioneer 11 anomalous acceleration (see Fig. 2) which, as reported by Nieto and Anderson (2005) [6], correspond to a time when the spacecraft was under the gravitational attraction of Jupiter and Saturn, are expected to be of a smaller magnitude and additionally, exhibit a large error due to the measurement of a fast changing quantity as they cross the gravitational field of the planets.

In regards to the Pioneer 11 Saturn flyby, we can rewrite (27) in terms of the planet's gravitational field:

$$\vec{a}_P = 2G\left(\frac{\sqrt{\rho}}{\sqrt{\rho'_{far}}} - 1\right)\left(\frac{M_S}{r_S^2}\frac{\vec{r}}{|r_S|} + \frac{M_{Sat}}{r_{Sat}^2}\frac{\vec{r}}{|r_{Sat}|}\right). \qquad (28)$$

Toth and Turyshev (2007) [20, p. 15] comment about the Pioneer 11's Saturn encounter:

> ...for Pioneer 11, a small value for the anomaly was found during the Jupiter–Saturn cruise phase. Right at the time of the Saturn encounter, however, when the spacecraft passed into a hyperbolic escape orbit, there was a rapid increase in the anomaly, whereafter it settled into the canonical value.

Unfortunately, no numerical values are quoted. However, in the light of Fig. 15 and (28) this text can be explained: When the spacecraft was in the incoming Saturn flyby, it went from an area of gravitational acceleration towards the Sun to an area of stronger gravitational acceleration towards Saturn. This has the effect of decreasing a_P until closest encounter.

Furthermore, as the spacecraft nears the planet it goes from the interstellar gravitational energy density (relatively low) and enters the area of Saturn's energy density with the highest value just at nearest encounter. In this area, $n' = \sqrt{\rho}\ /\ \sqrt{\rho'_{far}}$ increases towards a value similar to Earth's (see the value of n' for Saturn in Table I). Hence, the first term in brackets in (28) decreases rapidly until the nearest point to Jupiter and then increases rapidly settling in the interstellar n' value. This is precisely as reported by Toth and Turyshev.

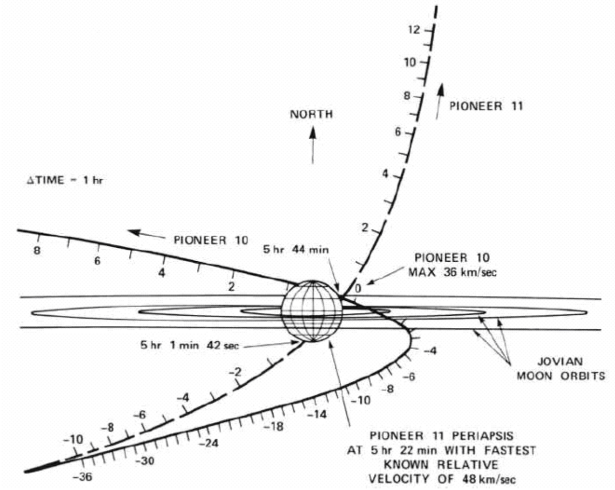


Fig. 13: The Jupiter Flyby of the Pioneer spacecraft, equatorial view [6, Fig. 2, p. 3].

5.4 Pioneer diurnal and annual Doppler residuals

In (16), the diurnal and annual residuals are essentially contained in the first bracket, namely

$$\left(1 + \frac{v_{ER}}{c}\cos\lambda\cos(\omega_R\, t + \phi_R) + \frac{v_{ET}}{c}\cos\lambda\cos(\omega_T\, t + \phi_T)\right)$$

which multiplies the "constant" term.

The term $\cos\lambda$ is the cosine of the DSN latitude. The latitude of the three stations are Goldstone = 35.4267° N, Madrid

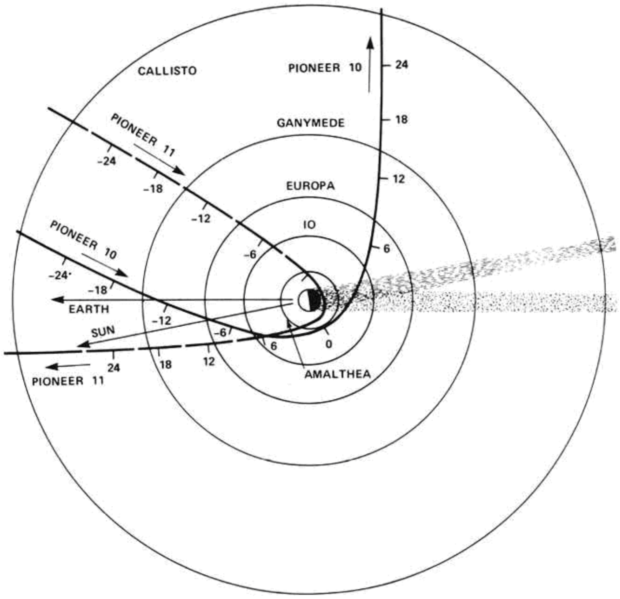


Fig. 14: The Jupiter Flyby of the Pioneer spacecraft. Polar view. From [6, Fig. 2, p. 3].

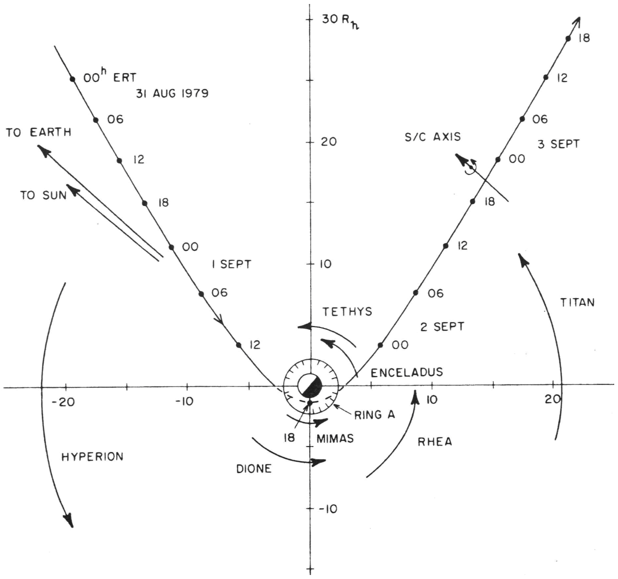


Fig. 15: Pioneer 11 Saturn flyby on the 1st of September 1979. From [6, Fig. 3, p. 6].

= 40.4314° N, and Canberra = 35.4013° S. We take the average 37.0864 so that the term $\cos\lambda = 0.79772$. If we substitute the Earth's equatorial rotational velocity $v_{ER} = 465.1$ m/s [30] and Earth's translational velocity $v_{ET} = 2.978 \times 10^4$ m/s and multiply by $\cos\lambda$, we get the following expression:

$$\Big(1 + 1.22402 \times 10^{-6} \cos(\omega_R\, t + \phi_R) + \\ + 7.92424386 \times 10^{-5} \cos(\omega_T\, t + \phi_T)\Big). \tag{29}$$

The "cos" term on the left describes the diurnal Doppler residuals with $\omega_R = 7.292 \times 10^{-5}$ rad/s, the Earth's sidereal angular rotation, and the "cos" term on the right describes the annual Doppler residuals with $\omega_T = 2.020$ x 10^{-7} rad/s, the Earth's sidereal angular rotation around the Sun. The sum of the three terms in (29) is very nearly equal to unity. For example, for the maximum amplitude of both oscillatory terms we get (29) to be: 1.00008046. On the other hand at opposition (Sun, Earth and Pioneer in that order) the third translational term is null and the maximum oscillatory term is due to Earth rotation and equal to 1.000001224.

The actual amplitude of the oscillations is obtained by multiplying (29) by the "constant" term in (16). As we have shown above, this is a variable value that depends on the distance to the Sun and also to the planets during encounters or flyby maneuvers. Hence to compare accurately with reported values, it is required to know at what distance or on what date were the measurements made.

5.4.1 Annual residuals

To compare with the Pioneer 10 sine wave reported, Fig. 4, we consider that for Pioneer 10 the data spans a distance from approximately 25 to 45 AU (as reported in Fig. 3). The result of multiplying the maximum amplitude of the oscillatory terms due to Earth's translation velocity, 1.00007924, by the calculated anomaly in this distance range, results in 5.1 to 1.6×10^{-8} cm/s^2. These values compare favorably to the amplitude of the annual oscillatory term reported by Anderson *et al* (2002) [3, p. 40] of "about 1.6×10^{-8} cm/s^2". However, they are larger than the estimate given by the same authors on [3, p. 14]: "the best estimate of the amplitude of the Pioneer 10 sine wave is $(0.525 \pm 0.155) \times 10^{-8}$ cm/s^2 and that of the Pioneer 11 wave is $(0.498 \pm 0.176) \times 10^{-8}$ cm/s^2. These values have a systematic error of $\sigma_{at} = 0.32 \times 10^{-8}$ cm/s^2" as reported for the entire Pioneer data span by Turyshev and Toth (2009) [24, p. 86].

The reported angular velocity of the annual Doppler residuals is approximately 2×10^{-7} rad/s. This value coincides with the Earth's sidereal translational rotation velocity which is 2.0200×10^{-7} rad/s as proposed in this work.

Figs. 5 and 6 show clearly the measured annual residuals. We wish to compare the results of the theory above to the undulating information contained in Fig. 5. To that end, we write the last term in (29) that deals with the annual Doppler residuals as a function of time for Pioneer 10 as:

$$(1 + 7.92424 \times 10^{-5} \cos(2.020 \times 10^{-7}\, t + \phi_R)) \tag{30}$$

where $\omega_T = 2.020 \times 10^{-7}$ rad/s is the Earth's angular rotation around the Sun.

Eq. (30) has to be multiplied by the calculated anomalous acceleration a_P. This value changes with distance as shown in Fig. 11. From 42 to 63 AU, the predicted anomaly calculated with (25) ranges from (1.84 to 0.837) $\times 10^{-8}$ cm/s^2 respectively. Hence we chose to select the middle of the three distance intervals as shown in Table IV (distance values derived using data in Fig. 3).

Interval	Dates	Mid date	Mid distance (AU)	Predicted a_P (10^{-8} cm/s^2)
I	1887-1990	1988.5	43	1.758
II	1990-1992	91	52	1.214
III	1993-1998	95.5	63	0.837

Table 4: Predicted a_P for the mid-distance of Pioneer 10 intervals. Values that were chosen to calculate the annual residuals.

Fig. 16 shows the agreement between the calculated annual Doppler residuals and the published results. The amplitude of the predicted oscillations are larger in interval I and decrease in intervals II and III as is reported by Anderson *et al* (2002): "At early times the annual term is largest. During Interval II, the interval of the large spinrate change anomaly, coherent oscillation is lost. During Interval III the oscillation is smaller and begins to die out." [3, p 40].

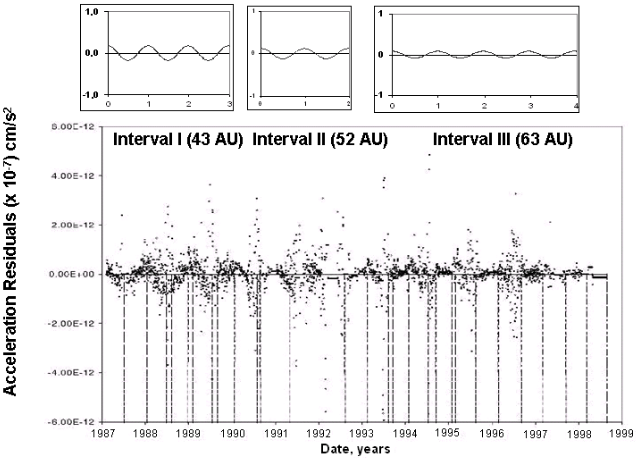


Fig. 16: Comparison of reported annual residual undulations with the predicted Doppler residuals. For uniformity, the original scale in units of [km/s^2] [3, Fig. 17, p. 40], is shown in units of [cm/s^2 $\times 10^{-7}$]. Inlays plots were drawn to approximately the same X–Y scale as the original data and show the predicted decreased calculated amplitudes corresponding to the center of each of the three intervals.

5.4.2 Diurnal residuals

Levi *et al* (2009), in their spectral analysis of the periodic terms yields the following measured periods: T1 = 0.9974 $\pm$ 0.0004 day, T2 = (1/2)(0.9972 $\pm$ 0.0004) day and T3 = 189 $\pm$ 32 days. As T = 0.9972 day = 1.0 sidereal day, these periods are consistent, within 0.02 percent, with variations of one sidereal day, half a sidereal day, and half a year.

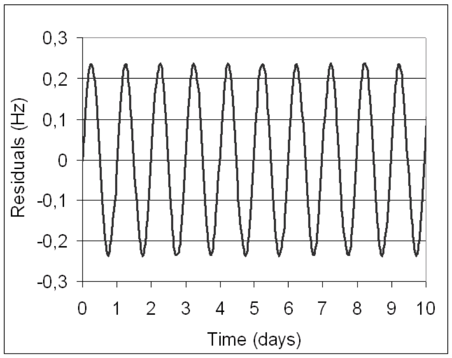


Fig. 17: Diurnal Doppler residuals calculated with (22).

The diurnal oscillations reported by Levi *et al* (2009) [9] between 23 November 1996 to 23 December 1996, reviewed above in Fig. 7 were at an estimated distance of 66.73 to 66.96 AU from the Sun (determined by the use of Fig. 3). Also, they were done at opposition, so that the annual rotational term is almost null and solar coronal influence is minimized.

The diurnal Doppler residuals in frequency units (Hz) may be calculated with (22a) namely:

$$\Delta f_D'' \approx 2f\,\frac{v_P\,v_{ER}}{c^2}\,\cos\lambda\cos\left(\omega_R\,t+\phi_R\right).$$

In this relation the speed of Pioneer v_P at a distance of 66.8 AU was estimated at 12 500 m/s and with the Earth's equatorial rotation velocity of 465.1 m/s, (22) leads to the oscillations shown in Fig. 17 next to the oscillations reported by Levi *et al* [9] in Fig. 18.

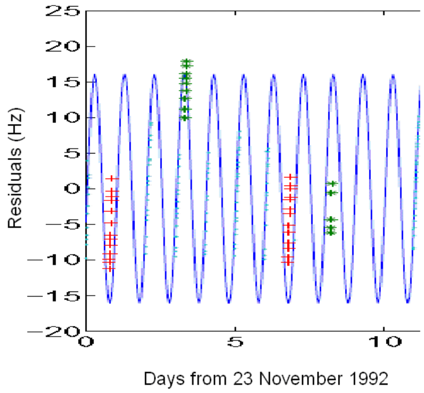


Fig. 18: Diurnal Doppler residuals reported by Levi *et al* [9].

We see that the frequency of diurnal oscillations reported by Levi *et al* (2009) [9] coincides with our predicted frequency $\omega_R = 2\pi f_E$, of Earth rotation, to within 0.02 percent.

A further conclusion of Levi *et al* (2009) [9, p. 10] is: "The main new result of the paper is that a large part of these diurnal and seasonal anomalies may be explained by a simple geometrical model where the light line on the tracking path is modified in a manner depending on the azimuthal angle ϕ between the Sun-Earth and Sun-probe lines.”

We reflect about this conclusion that the azimuthal angle ϕ between the Sun-Earth and Sun-probe lines will show diurnal variations superimposed on annual variations which are wholly compatible with the first bracket of (16) and expression (29) above. With the use of (24a), namely:

$$\Delta v_D'' = \frac{v_P \, v_{ER}}{c} \cos\lambda \cos(\omega_R \, t + \phi_R)\,,$$

we can calculate the diurnal Doppler residuals in velocity units as reported by Anderson *et al* (2002) [3] and shown in Fig. 4, using the rotational velocity of Earth 465,1 m/s, and the estimated speed of Pioneer 10 in 1995 of 12 500 m/s. A comparison of the results is shown in Fig. 19 and Fig. 20.

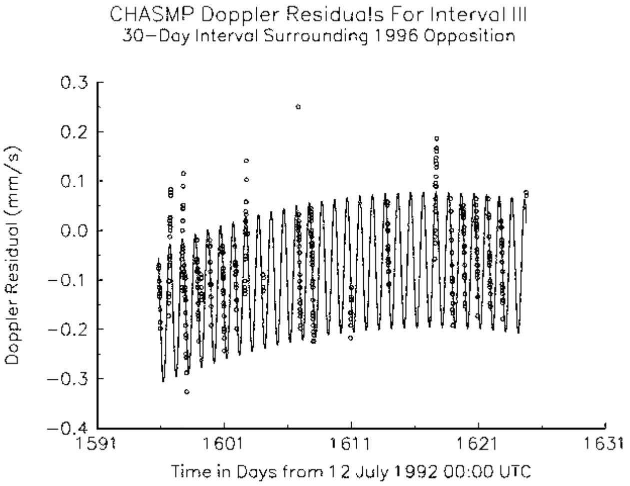


Fig. 19: Diurnal Doppler residuals in velocity units from [3, Fig. 18, p. 41].

5.5 On the energy density due to far away stars and galaxies

In this work, we have used the value of $\rho^* = 1.094291 \times 10^{15}$ J/m^3, the energy density of space due to far away stars and galaxies, a value calculated by J. Céspedes-Curé [26, p 279], obtained using starlight deflection measurements during total sun eclipses. With this value in the equations, in this work, it has been possible to calculate numerically the anomalous Pioneer acceleration.

It is possible to work in the inverse order and use the empirically determined values of the anomaly to calculate in an independent way the value of this physical magnitude. When this is done, using the accurately reported Pioneer Anomaly at 20 AU, the result gives for the energy density of space due to far away stars and galaxies the value $\rho^* = 1.0838 \times 10^{15}$ J/m^3.

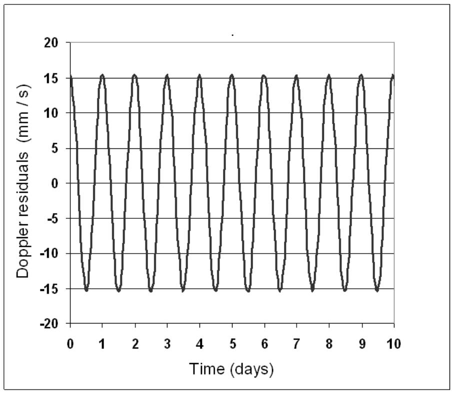


Fig. 20: Diurnal Doppler residuals in velocity units calculated with the use of (24a).

This value differs by less than 1 percent from the value determined by J. Céspedes-Curé on the basis of a completely different phenomenon, the bending of light rays during solar eclipses.

We would like to briefly review the procedure that was published to make this determination. For details please consult [4]. The calculation uses the following formulas: Eq. (19) in [4]:

$$n' = 1 - \frac{E_D \, c}{2 f_e \, G \left(\frac{M_S}{r_S^2} + \frac{M_E}{r_E^2} \right)}\,, \tag{31}$$

and Eq. (8) in [4]:

$$\rho^* = \frac{\rho_{\text{Sfar}} + \rho_{\text{Efar}} - n'^2(\rho_{S\,1AU} + \rho_E)}{n'^2 - 1} \tag{32}$$

where (numerical values in SI units)

- n' = index of refraction of space at 20 AU (comes out to 0.999973567943846).
- ρ^* = energy density of space due to far away stars and galaxies.
- E_D = a steady frequency drift of 5.99×10^{-9} Hz/s from the Pioneer 10 spacecraft [3, p. 20].
- f_e = 2 295 MHz, the frequency used in the transmission to the pioneer spacecraft [3, p. 15].
- c = 299792458.0 m/s. Speed of light on Earth at the surface.
- $G = 6.67300 \times 10^{-11}$ m^3 kg^{-1} s^{-2}, Newton’s universal constant of gravitation.
- $M_S = 1.98892 \times 10^{30}$ kg, mass of the Sun.
- $M_E = 5.976 \times 10^{24}$ kg, mass of the Earth.
- The distances r_S and r_E are the distances from the spacecraft at 20 AU (20 AU from the Sun, 19 AU from the Earth) to the center of the Sun and Earth respectively. To calculate them use was made of:

- 1 Astronomical Unit (AU) = 149 598 000 000 meters.

To calculate Eq. (8) in [4], use is made of the energy density given by our Eq. (2), namely

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

where r is the distance from the centre of the Sun or Earth to the point where the energy density is being calculated as follows:

- for the Earth surface: $r_E = 6\,378\,140$ m, radius of the Earth.
- for the Sun at 1 AU: $r_S = 149\,598\,000\,000$ meters.
- for the Sun at 20 AU: twenty times the previous value used to calculate ρ_{Sfar}.
- for the Earth at 20 AU: radius of the Earth + 19 times 149 598 000 000 meters used to calculate ρ_{Efar}.

Values were calculated with Microsoft Office Excel 2003 which uses 15 significant digits.

6 Discussion

The theoretically calculated Pioneer Anomalous acceleration shown in Fig. 11 has a decreasing value as a function of distance contrary to the generally accepted opinion that it is a "constant" value. However, the numerical evidence supplied by V. Toth (2009) [22] in his independent analysis, gives confirmation that the anomaly is a decreasing function which coincides exactly with the theoretical slope for Pioneer 11 at a distance between 29 and 38 AU and also with the theoretical slope for Pioneer 10 at a distance from the Sun between 42 and 48 AU.

At a distance from the Sun of 20 AU, the theoretical curve predicts $a_P = 7.754{\times}10^{-8}$ cm/s^2 which differs by just 1.2 percent from the value $a_P = 7.85 \pm 0.02 \times 10^{-8}$ cm/s^2 reported by O. Olsen (2007) [7] in his independent analysis of the Pioneer 10 data for Interval I. Furthermore, the theoretical value differs by just −0.2 percent from the 1-day batch-sequential estimation for the entire data interval of 11.5 years for Pioneer 10 reported by Anderson *et al* (2002) [3, p. 24] .

The theory predicts that the anomalous acceleration has a vectorial character $\vec{a}_P$ in the direction of the resultant gravitational acceleration field at the position of the spacecraft. This fact allows satisfactory explanation of the reported anomalous behavior of Ulysses, Pioneer 10 and 11 during Jupiter flyby. The observations of the peculiar values reported for the first 3 values of Pioneer 11 (see Fig. 11) are adequately explained with consideration that they correspond to the spacecraft being affected by the Jupiter gravitational acceleration which at close distances exceeds the Sun's gravitational acceleration (see Fig. 12). The prediction that the anomalous acceleration is in the direction of the resultant gravitational acceleration field at the position of the spacecraft gives an answer to this question, which is posed by several publications on the Pioneer Anomaly.

With regard to the harmonic behavior of the Doppler residuals, relaxing the assumption that the value of the speed of light c in the Doppler formula is invariant and adopting the galilean addition of the Earth rotational and translational velocity to the speed of light, results in an almost exact agreement with the measured *frequencies* for the annual (within 1 percent) and diurnal (within 0.02 percent) residuals as shown in Sections 5.4.1 and 5.4.2 above. However, the values for the magnitudes of the oscillations do not all agree as well.

In the case of the annual residuals, we do a visual comparison in Fig. 16 which agrees quite well. Particularly if we take into account that the reported values have significant errors: A systematic error of $\sigma_{at} = 0.32 \times 10^{-8}$ cm/s^2 ($\sigma_{at} = 3.2 \times 10^{-7}$ cm/s^2 in the scale of Fig. 16) is reported for the entire Pioneer data span by Turyshev and Toth (2009) [24, p. 86]. Considering the scatter of the measured values, the predicted magnitude adequately fits the data in this case.

In the case of the diurnal residuals, expressed in frequency units (Hz) as shown in Fig. 13, there is a discrepancy in the *amplitude* of the order of a factor of about 70 smaller in the calculated value of the oscillations in comparison with the amplitude of the oscillations reported by Levi *et al* [9]. With the calculated oscillations in velocity units (mm/s) the reverse is obtained. As shown in Fig. 19 the calculated amplitude is a factor of about 50 larger than the values in Fig. 4 by Anderson *et al* [3]. In view of these differences it is instructive to compare the amplitudes of the different reported values which also show significant differences.

The amplitude of the diurnal residuals in *frequency* units (Hz) reported by Levi *et al* [9, p. 6], shown in Fig. 7 are a factor of about 250 times greater than the amplitude of diurnal residuals in frequency units (Hz) reported by Anderson *et al* (2002), our Fig. 4 [3, Fig. 18, p. 41]. Both reports are for the same interval of time (23 November to 23 December 1996).

Regarding the annual residuals there is also a discrepancy in the amplitudes expressed in *acceleration* units (cm/s^2) reported by Anderson *et al* in the 2002 paper. The amplitude of the annual oscillations shown in Fig. 4 are about 10 times greater than those reported in the same paper in Fig. 12 [3, p. 26]. In view of the good agreement in the prediction of the frequencies of the harmonic Doppler residuals, it is not clear what are the sources of the discrepancies between reported amplitudes, or the differences between reported and the calculated amplitudes.

7 Conclusions

As summarized in the Discussion above, the theory presented in this work is capable of explaining qualitatively and quantitatively the phenomena associated with the Pioneer Anomaly, both, the secular and the harmonic terms that up to now had no plausible explanation. These precise calculations of the

Pioneer Anomaly, without any adjustable parameters, provide additional confirmation of the Céspedes-Curé hypothesis, that c the speed of light depends on the gravitational energy density of space as defined by (1) namely: $c = k / \sqrt{\rho}$. The highly accurate calculation of the observed frequencies of the annual and diurnal Doppler residuals and some of the amplitudes of the annual oscillations supply additional evidence that the speed of Earth adds to c, the speed of light, according to the galilean addition of velocity, thereby confirming this conclusion put forth by the analysis of S. Gift using the Global Positioning System [16–18].

The extremely accurate measurements provided by NASA as empirical data and the theoretical explanation, agreeing within 1 percent, presented in this work for the Céspedes-Curé hypothesis, have profound consequences in the current cosmology theories. The dependence of the speed of light on the gravitational energy density of space implies a revision of all astronomical measurements of velocity based on the Doppler, blue and red shifts, of stars and galaxies. These have importance in the determination of matters such as the Hubble constant, the expansion of the universe, the flat rotation curve of galaxies (which gave birth to the theory of dark matter) and the extreme values of the redshifts of very far away galaxies (so called inflation) which gave birth to the theory of dark energy.

Acknowledgements

We would like to acknowledge the independent verification of numerical calculations provided by Simon E. Greaves.

Received on June 3, 2021

References

1. Anderson J. D., Campbell J. K., Ekelund J. E., Jordan E. and Jordan J. F. Anomalous Orbital-Energy Changes Observed during Spacecraft Flybys of Earth. *Phys. Rev. Letters*, 2008, v. 100 (091102), 1–4.
2. Greaves E. D., Bracho C. and Mikoss I. A Solution to the Flyby Anomaly Riddle. *Progress in Physics*, 2020, v. 16 (1), 49–57.
3. Anderson J. D., Laing Ph. A., Lau E. L., Liu A. S., Nieto M. M. and Turyshev S. G. Study of the anomalous acceleration of Pioneer 10 and 11. *Phys. Rev. D*, 2002, v. 65 (082004).
4. Greaves E. D. NASAs astonishing evidence that c is not constant: The pioneer anomaly. arXiv: physics.gen-ph/0701130v1.
5. Turyshev S. G. and Toth V. T. The Pioneer Anomaly. arXiv: gr-qc/1001.3686v2.
6. Nieto M. M. and Anderson J. D. Using early data to illuminate the Pioneer anomaly. *Classical and Quantum Gravity*, 2005, v. 22, 5343–5354. arXiv: gr-qc/0507052v2.
7. Olsen O. The constancy of the Pioneer anomalous acceleration. *Astronomy and Astrophysics*, 2007, v. 463, 393–397.
8. Ghosh A. On the Annual and Diurnal Variations of the Anomalous Acceleration of Pioneer 10. *Apeiron*, 2007, v. 14 (3).
9. Levy A., Christophe B., Berio P., Metris G., Courty J. M. and Reynaud S. Pioneer Doppler data analysis: study of periodic anomalies. *Adv. Space Res.*, 2009, v. 43, 1538–1544. arXiv gr-qc/0809.2682.
10. Turyshev S. G., Toth V. T., Kinsella G., Lee S.-C., Lok S. M., Ellis, J. Support for the Thermal Origin of the Pioneer Anomaly. *Physical Review Letters*, 2012, v. 108 (24), 241101. arXiv: /1204.2507.
11. Bilbao L. Does the velocity of light depend on the source movement? *Prog. in Phys.*, 2016, v, 12 (4), 307–312.
12. Rievers B. and Lämmerzahl C. High precision thermal modeling of complex systems with application to the Flyby and Pioneer Anomaly. *Annalen der Physik*, 2011, v, 523 (6), 439–449.
13. Francisco F., Bertolami O., Gil P. J. S. and Páramos J. Modelling the reflective thermal contribution to the acceleration of the Pioneer spacecraft. arXiv: physics.space-ph/1103.5222v2.
14. Bilbao L., Bernal L. and Minotti F. Vibrating Rays Theory. arxiv: abs/1407.5001.
15. Greaves E. D. A Neo-Newtonian Explanation of the Pioneer Anomaly. *Revista Mexicana de Astronomia y Astrofisica*, 2009, v. 35, 23–24. ISSN 0185-1101, www.redalyc.org/articulo.oa?id=57115758008, accessed August 2021.
16. Gift S. J. G. Doppler Shift Reveals Light Speed Variation. *Apeiron*, 2010, v. 17 (1), 13–21.
17. Gift S. J. G. Time Transfer and the Sagnac Correction in the GPS. *Applied Physics Research*, 2014, v. 6 (6), 1–9.
18. Gift S, J. G. One-way Speed of Light Using the Global Positioning System. In: Torres, G., ed. Global Positioning Systems (GPS): Performance, Challenges and Emerging Technologies, Nova Publishers, NY, 2017, pp. 45–66.
19. NASA Solar System Exploration – Galileo. solarsystem.nasa.gov/missions/galileo/in-depth/, retrieved August 2020.
20. Toth V. T. and Turyshev S. G. The Pioneer Anomaly: Seeking an explanation in newly recovered data. *Can. J. Phys.*, 2007, v. 84, 1063–1087.
21. Markwardt C. Independent Confirmation of the Pioneer 10 Anomalous Acceleration. arXiv: gr-qc/0208046v1.
22. Toth V. T. Independent Analysis of the Orbits of Pioneer 10 and 11. *Int. J. Mod. Phys.*, 2009, v. D18, 717–741. arXiv: /0901.3466.
23. Levy A., Christophe B., Bério P., Métris G., Courty J.-M. and Reynaud S. Pioneer 10 Doppler data analysis: disentangling periodic and secular anomalies. *Advances in Space Research*, 2009, v. 43, 1538–1544. arXiv: gr-qc/0809.2682v2,
24. Turyshev S. G. and Toth V. T. The Pioneer Anomaly in the Light of New Data. arXiv: gr-qc/0906.0399v1.
25. Pioneer 10 coordinates: theskylive.com/pioneer10-info. Pioneer 11 coordinates: theskylive.com/pioneer11-info. Topocentric coordinates computed for the selected location of Greenwich, United Kingdom, accessed July 2021.
26. Céspedes-Curé J. Einstein on Trial or Metaphysical Principles of Natural Philosophy. et al. Organization, Caracas, Venezuela, 2002. www.nuclear.fis.usb.ve/Cespedes-Cure-2002-Einstein-on-Trial-J.pdf, retrieved 10 January 2019.
27. Greaves E. D. (2015) La hipótesis de Céspedes-Curé y el índice de refracción del espacio en un campo magnético. (The Céspedes-Curé hypothesis and the index of refraction in a magnetic field). *Acta Científica Venezolana*, 2015, v. 66 (4), 226–229.
28. Greaves E. D. The index of refraction of quasi-empty space. Unpublished, Universidad Simón Bolívar, Caracas, Venezuela, 2015. www.nuclear.fis.usb.ve/fn/wp-content/uploads/2015/07/GREAVES-ED-Index-of-refraction-of-quasi-empty-space-V11.pdf, retrieved 19 April 2019.
29. Greaves E. D. Propiedades del espacio vacío. (Properties of empty space). Memorias del II Congreso de ABAE, September 18–22, 2017.
30. Cox, A. N. ed. Allen's Astrophysical Quantities. Springer, 2015, p. 244.

Gamow Theory for Transmission and Decay of Unbound Diprotons

Tianxi Zhang

Department of Physics, Chemistry, and Mathematics, Alabama A & M University, Normal, Alabama 35762, USA.
E-mail: tianxi.zhang@aamu.edu

Transmission and decay of unbound diprotons have been investigated in accordance with the Gamow theory for the quantum tunneling and radioactive decays. It is shown that a diproton, once formed, will be quickly decayed with two typical decay modes: (1) the proton decay, which causes the diproton to be separated into two separate protons and (2) the β^+ decay, which causes the diproton to be changed and fused into a deuteron after emitting a positron and a neutrino. For both of the decay modes, the transmission probabilities rapidly increase with the energy of the emitted particle. The β^+ decay from a diproton is much rarer ($< 10^{-4}$ times less) in general than the proton decay. The lifetimes for both of the two decay modes slowly decrease with the energy of the emitted particle and are extremely short to about 10^{-21} s. In addition, we have also modeled the diproton decay of a typical proton-rich radioactive heavy nucleus such as ^{15}Ne and obtained result of lifetimes consistent with measurements

1 Introduction

Helium-2 or ^{2}He is an isotope of helium. Its nucleus consists of only two protons and is usually called a diproton. It is extremely unstable and believed to be in an unbound state with a negative binding energy due to the spins of the two protons to be anti-aligned according to the Pauli Exclusion Principle [1, 2]. A diproton can be formed in two ways: (1) by combination of two separate protons or (2) by decay from radioactive heavy nuclei. Two separate protons, when they collide with enough energy to tunnel through the Coulomb barrier between them, form a diproton, $^1\text{H}+^1\text{H}+\text{Energy} \longrightarrow {}^2\text{He}$. On the other hand, some proton-rich (or neutron-rare) heavy nuclei have been experimentally found to emit diprotons. For instances, the radioactive nuclei ^{15}Ne and ^{11}O can decay, respectively, to ^{13}O and ^{9}C after emitting a diproton [3, 4]. This type of event for a diproton to be emitted from a radioactive nucleus is usually called the diproton decay.

A diproton, once formed via either one of the two ways as described above, will quickly decay through either one of the two different modes [5]. It most likely undergoes a proton decay to change immediately back to two separated protons, $^2\text{He} \longrightarrow {}^1\text{H} + {}^1\text{H}$, with a probability greater than 99.99%. In this case, both of the emitted particle and the leftover nucleus are protons. The formed diproton can also very rarely undergo a positron (or β^+) decay and get fused to form a deuteron, $^2\text{He} \longrightarrow {}^2\text{H} + \text{e}^+ + \nu_\text{e}$, with a probability less than 0.01%. In this case, one of the two protons in the formed diproton decays to a neutron after emitting a positron and a neutrino. Meanwhile, the neutron immediately fuses with the other proton to form a deuteron and release nuclear energy. It can be seen that the β^+ decay of diprotons is much rarer (about ten thousand or more times rarer) than the proton decay of diprotons. The lifetime of a diproton is extremely short and believed to be much much less than 10^{-9} s. Up to now, scientists have only provided these upper bound values for both of the rareness of β^+ decay and the lifetime of diprotons. The actual rareness of the β^+ decay and the lifetime of diprotons are still uncertain.

The Sun is a giant natural fusion reactor with an emission power of 3.85×10^{26} W from the nuclear fusion of its core's 1.2×10^{56} protons at a rate of about 3.6×10^{38} protons per second to produce helium nuclei or α-particles [6]. A diproton is an intermediate in the first step of the proton-proton chain nuclear reaction that is occurring in the cores of stars including our Sun. Therefore, the instability of diprotons critically affects the rate of nuclear fusion reactions in the core of the Sun. From classical physics, no proton should be able to overcome the 820 keV Coulomb barrier between protons to form a diproton and then get fused in the Sun's core, where the temperature is about 1.5 keV. According to Gamow's theory or model for the quantum tunneling probability [7], however, one part per million of the core's protons can penetrate or tunnel through the Coulomb barrier to form diprotons. Considering the high ion-collision frequency (over about 20 terahertz), one can find approximately 10^{63} sufficient collisions for diprotons to be formed in one second in the core of the Sun. Even though as mentioned above less than 0.01% of diprotons are fused to deuterons via the β^+ decay, the fusion reaction rate in the core of the Sun is still around 10^{21} times higher in magnitude than the actually observed fusion reaction (or power emission) rate. This extremely high fusion rate would lead the Sun to have an intensive explosion, if there does not exist any other fusion inhibitors.

Recently, the author proposed that the plasma waves, globally destabilized in the core of the Sun, can significantly reduce the nuclear fusion reaction rate to the observed power emission rate or luminosity and thus effectively prevent the Sun from an instantaneous explosion [8]. Through significantly reducing the electric permittivity of the core plasma, plasma waves can extremely raise the Coulomb barrier and shift the Gamow peak to a higher energy of particles to extremely inhibit the fusion reaction. It has been shown that, if the frequency of plasma waves that are globally generated in the core plasma of turbulences is about 1.28 times the plasma frequency, the Sun can have the actual fusion rate or shine on at the currently observed luminosity. This implies that, in addition to the quantum tunneling effect and rareness of β^+ decay, plasma waves are also playing the essential role in solar nuclear fusion and power emission.

In this paper, we study the transmission and lifetime for the proton and β^+ decays of unbound diprotons according to the Gamow theory for the quantum tunneling. We obtain that the transmission probability and lifetime of unbound diprotons depend on the energy of the emitted or decayed particles. When the energy of emitted protons is about 800 keV or higher, more than 99.99% of diprotons will decay into separate protons. When the energy of emitted positrons is about 10 eV or lower, less than 0.01% of diprotons will decay and fuse to deuterons. The lifetimes of a diproton via both of the two decay modes decrease with the energy of emitted particles and are about 10^{-21} s or shorter. The speeds of a proton with hundreds of keV and an electrons with several eV are typically valued at about 10^6 m/s.

2 Gamow theory for transmission and decay of diprotons

In 1928, George Gamow proposed a theory for α-decay of radioactive heavy nuclei [7]. Since the α particle, i.e. the helium nucleus, is a positively charged particle (with charge $Z_1 e$, where $Z_1 = 2$ for the α particle), it will be electrically repelled by and further escape from the leftover nucleus (with charge $Z_2 e$). Here Z_1 and Z_2 are the atomic numbers of the nuclear elements or the proton number in the nucleus of the emitted particle and the leftover nucleus, $\epsilon_0 = 8.85 \times 10^{-12}$ $C^2/(J\,m)$ is the permittivity of free space, and $e = 1.6 \times 10^{-19}$ C is the charge of the proton. Gamow's theory approximately modeled the potential energy by a finite potential square well to represent the attractive nuclear force and joined with a Coulomb repulsive potential tail [9],

$$V(r) = \begin{cases} -V_0 & \text{for } 0 < r < r_1 \\ \frac{1}{4\pi\epsilon_0}\frac{Z_1 Z_2 e^2}{r} & \text{for } r_1 < r < \infty \end{cases} . \tag{1}$$

Fig. 1 sketches the potential energy $V(r)$ given by (1) as a function of radial distance r in all the classical and quantum regions. The width of the potential square well is noted by r_1, which is determined by the radius of the nucleus or by the sum of the radii of both the emitted particle and the leftover nucleus. The depth of the potential square well is noted by V_0, which is much greater than the maximum height of the Coulomb barrier, U_c. The outer turning point (i.e. r_2) can be determined, in terms of the energy E of the emitted α particle to be equal to the potential energy at r_2, by

$$r_2 = \frac{4\pi\epsilon_0 E}{Z_1 Z_2 e^2} . \tag{2}$$

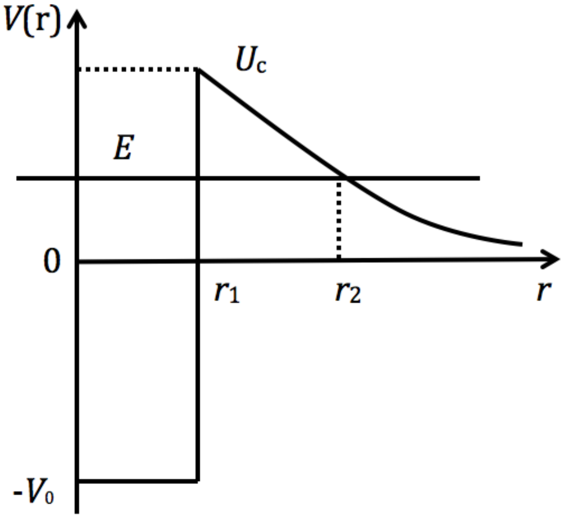


Fig. 1: Gamow's modeling of the potential energy for an electrically charged particle to decay or be emitted from a radioactive nucleus. It consists of the potential energy square well for the attractive nuclear force and the Coulomb potential energy tail for the repulsive electric force between the emitted particle and the leftover nucleus of the decay.

In the central potential $V(r)$, the radial Schrödinger equation is,

$$\frac{d^2 u(r)}{dr^2} = \frac{2\mu}{\hbar^2}[V(r) - E]u(r) + \frac{l(l+1)}{r^2}u(r) , \tag{3}$$

where $u(r)$ is the radial wave function, μ is the reduced mass, $\mu = m_1 m_2/(m_1 + m_2)$ with m_1 the mass of the emitted particle and m_2 the mass of the leftover nucleus. The integer l is the quantum number for the magnitude of angular momentum and $\hbar$ is defined by $\hbar = h/2\pi$ with $h = 6.62 \times 10^{-34}$ J s, the Planck constant. A two-body system with a central force or potential can be treated as a system of one body with the reduced mass.

Applying the WKB approximation and considering the case of $l = 0$, one can approximately solve the radial Schrödinger equation and find the radial wave functions to be

$$u(r) = \frac{C}{\sqrt{|p(r)|}} \exp\left[\pm\frac{1}{\hbar}\int |p(r)| dr\right] , \tag{4}$$

where $p(r)$ is defined by

$$p(r) = \sqrt{E - V(r)} . \tag{5}$$

Here it should be pointed out that the general solution of the radial Schrödinger equation should be the combination of these two.

Then, from the solved wave function, the transmission (or tunneling) probability is obtained as

$$T = e^{-2\gamma} , \tag{6}$$

where γ is determined by

$$\gamma = \frac{1}{\hbar}\int_{r_1}^{r_2} dr\,\sqrt{E - V(r)} = \frac{\sqrt{2\mu E}}{\hbar}\left[r_2\left(\frac{\pi}{2} - \arcsin\sqrt{\frac{r_1}{r_2}}\right) - \sqrt{r_1(r_2 - r_1)}\right]. \tag{7}$$

And the lifetime of the parent nucleus is given by

$$\tau = \frac{2r_1}{v}e^{2\gamma} \tag{8}$$

where $v = \sqrt{2E/m_1}$ is the speed of the emitted (or α) particle. It should be noted that, although being proposed for explaining the α decay of radioactive nuclei, the Gamow model is applicable in general for the decay or emission of any type of charged particles from a radioactive nucleus such as the proton decay from a diproton, β^+ decay from a diproton, and emission of a diproton from a radioactive heavy nucleus (e.g. diproton decays of ^{15}Ne and ^{11}O), and so on.

For the proton decay mode of a diproton, the emitted particle is a proton and the leftover nucleus is also a proton. In this case, we have $Z_1 = Z_2 = 1$, $m_1 = m_2 = m_p$, and $\mu = m_p/2$, where $m_p = 1.67 \times 10^{-27}$ kg is the proton mass. The width of the potential square well or the radius of the diproton can be chosen as $r_1 = 1.75 \times 10^{-15}$ m. With the values of these parameters and (6)–(8), we can plot, in Fig. 2, the transmission probability for the proton decay of the diproton (solid line) and the lifetime of the diproton via the proton decay mode (dashed line) as a function of the energy of the proton. It is seen that the transmission probability increases with the energy. Most diprotons undergo this decay mode when the energy of the emitted particle is greater than about some hundred keV. In other words, diprotons rarely decay into protons with energy much below about the Coulomb barrier such as one hundred keV or less. The lifetime of unbound diprotons via this decay mode is very short and slowly decreases with the energy of the emitted particle. When the energy of the emitted particle is greater than about some hundred keV, the lifetime of diprotons is as short as about 10^{-21} s.

For the β^+ decay mode of a diproton, the emitted particle is a positron and the leftover nucleus is a deuteron. In this case, we have $Z_1 = Z_2 = 1$, $m_1 = m_e$, $m_2 = 2m_p$, $\mu = m_e$, where $m_e = 9.1 \times 10^{-31}$ kg is the electron mass. The width of the potential square well or the radius of the diproton can be chosen again as $r_1 = 1.75 \times 10^{-15}$ m. With the values of these parameters and (6)–(8), we can plot, in Fig. 3, the transmission probability for the β^+ decay of a diproton (solid line) and the lifetime of diproton via this decay mode (dashed line) as a function of the energy of the positron. It is seen that the transmission probability increases with the energy. Diprotons rarely undergo this decay mode when the energy of the positron is less than about some hundred eV. The reason for the β^+ decay of the diproton to be extremely rare is because the energy of the emitted positron is far below the 820 keV Coulomb barrier. For the transmission probability to be about 10^{-21}, the energy of the emitted positron must be less than an eV, which may not be reasonable. Therefore, the result obtained here supports the existence of other physics effects such as plasma oscillations or waves that the author recently proposed to significantly inhibit the nuclear fusion reaction in the core of the Sun [8]. The lifetime of unbound diprotons via this β^+ decay mode is also very short and slowly decreases with the energy of the emitted positron. When the energy of the emitted positron is as high as about some hundred eV, the lifetime of diprotons is also as short as about 10^{-21} s.

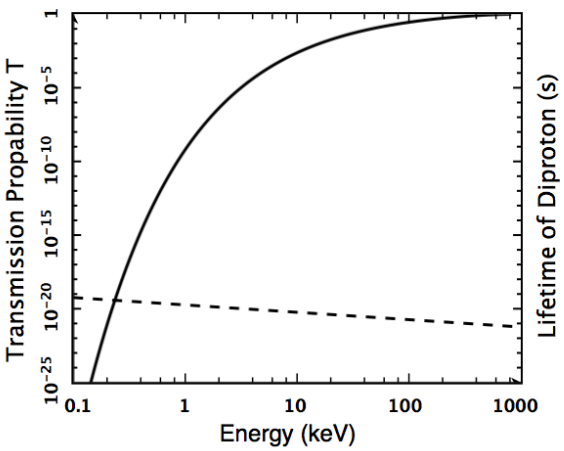


Fig. 2: Proton decay and lifetime of an unbound diproton. The solid line plots the transmission probability of a proton from the unbound diproton in the potential energy well to tunnel through the Coulomb barrier as a function of the energy of the proton. The dashed line plots the lifetime of the diproton.

For the diproton decay of radioactive heavy nuclei such as ^{15}Ne, the emitted particle is a diproton and the leftover nucleus is ^{13}O. In this case, we have $Z_1 = 2$, $Z_2 = 8$, $m_1 = 2m_p$, $m_2 = 13m_p$, $\mu = 1.73\,m_p$, Here we have considered approximately both proton and neutron having about the same mass. The width of the potential square well or the radius of ^{15}Ne nucleus can be chosen as $r_1 = 4 \times 10^{-15}$ m. With the values of these parameters and (6–8), we can plot, in Fig. 4, the transmission probability for the diproton decay from a radioactive nucleus ^{15}Ne (solid line) and the lifetime of the nucleus ^{15}Ne via this diproton decay mode (dashed line) as a function of the energy of the diproton. It is seen that the transmission probability increases with the energy. Most ^{15}Ne nuclei undergo the diproton decay when the energy of the emitted particle is greater than about some MeV. The lifetime of the radioactive nucleus ^{15}Ne via the diproton decay mode is very short and slowly decreases with the energy of the emitted diproton. When the energy of the emitted diproton is as high as about some MeV, the lifetime of the radioactive nucleus ^{15}Ne is

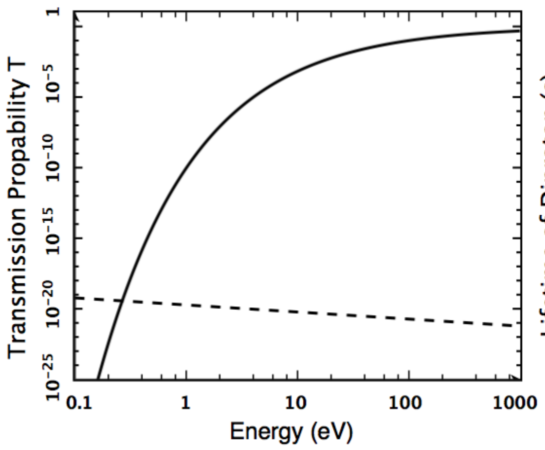


Fig. 3: Positron (or β^+) decay and lifetime of an unbound diproton. The solid line plots the transmission probability of a positron from the unbound diproton in the potential energy well to tunnel through the Coulomb barrier as a function of the energy of the positron. The dashed line plots the lifetime of the diproton.

as short as about 10^{-21} s, consistent with measurements [10]. The diproton decay was also detected from other nuclei such as ^{18}Ne nucleus [11, 12].

3 Discussions and Conclusions

If diprotons are bound, stars would burn about a billion billion times brighter in luminosity or faster in nuclear reaction, resulting in a universe to fail the life support [13, 14]. This diproton disaster can be overcome by plasma oscillations or waves, which have been shown recently to be able to be extremely efficient in inhibiting the nuclear reaction [8], to have the observed luminosity without need to adjust the stars' central temperature, density, and initial number of deuterons. In future study, we will study in more detail the transmission probability of bound diprotons for the fusion reaction.

As a consequence of this study, we have investigated the transmission and decay of unbound diprotons according to the Gamow theory. An unbound diproton is extremely unstable and quickly decays through two types of decay modes with lifetime to be extremely short down to about 10^{-21} s and transmission probability to be significantly energy dependent. A diproton mostly undergoes a proton decay to be two separate protons with a transmission probability higher than 99.99%, and rarely undergoes a β^+ decay to form a deuteron with a transmission probability lower than 0.01%. In the reasonable energy range, the β^+ decay of diproton is not rare enough for the Sun to have the observed reaction rate, which supports the author's recently proposed other inhibition effect such as plasma oscillation in solar nuclear fusion. The result obtained for the diproton decay from a radioactive nucleus can also be consistent with measurements.

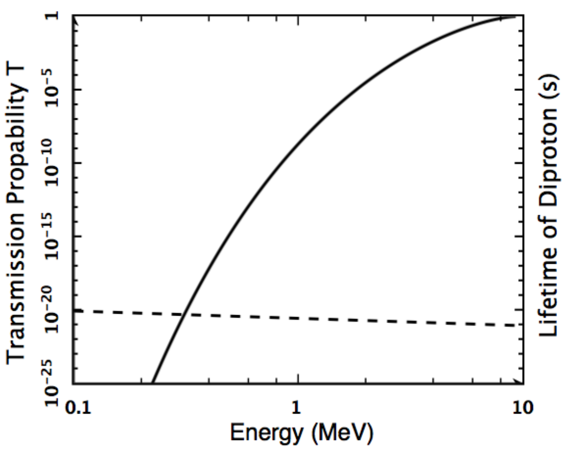


Fig. 4: Diproton decay and lifetime of ^{15}Ne nucleus. The solid line plots the transmission probability of a diproton from the radioactive ^{15}Ne nucleus in the potential energy well to tunnel through the Coulomb barrier as a function of the energy of the diproton. The dashed line plots the lifetime of ^{15}Ne nucleus.

Acknowledgements

The author acknowledges the reviewer and the editor for improving the manuscript quality.

Received on June 9, 2021

References

1. Pauli W. Exclusion Principle and Quantum Mechanics. Writings on Physics and Philosophy, 1946, 165–1981.
2. Bertulani C. A. Nuclear Physics in a Nutshell. Princeton University Press, 2007, ISBN 978-0-691-12505-3.
3. Wamers F. *et al.* First Observation of the Unbound Nucleus ^{15}Ne. *Physical Review letters*, 2014, v. 112, 132502.
4. Web T. B. *et al.* First Observation of Unbound ^{11}O, the Minor of the Halo Nucleus ^{11}Li. *Physical Review Letters*, 2019, v. 122, 122501.
5. https://en.wikipedia.org/wiki/Isotopes of helium.
6. Zirin H. Astrophysics of the Sun. Cambridge and New York, Cambridge Univ. Press, 1988.
7. Gamow G. Zur Quantentheorie des Atomkernes. *Z. Physik*, 1928, v. 51, 204–212.
8. Zhang T. X. The Role of Plasma Oscillation Played in Solar Nuclear Fusion. *Progress in Physics*, 2021, v. 17, 67–71.
9. Grffiths D. J. Introduction to Quantum Mechanics, 2nd Edition. Person Prentice Hall, 2005.
10. Blank B., Ploszajczak M. Tow-Proton Radioactivity. *Reports on Progress in Physics*, 2008, v. 71, 046301.
11. del Campo G. J. *et al.* Decay of a Resonance in ^{18}Ne by the Simultaneous Emission of Two Protons. *Physical Review Letters*, 2001, v. 66, 43–46.
12. Raciti G. *et al.* Experimental Evidence of ^{2}He Decay from ^{18}Ne Excited States. *Physical Review Letters*, 2008, v. 100, 192503–192505.
13. Bradford R. A. W. The Effect of Hypothetical Diproton Stability on the Universe. *J. of Astrophys. Astron.*, 2009, v. 30, 119–131.
14. Barnes L. A. Binding the Diproton in Stars: Anthropic Limits on the Strength of Gravity. *JCAP*, 2015, No. 12, 050.

Entropy Analysis of the Bioelectrical Activity of Plants

Hartmut Müller, Roberta Baccara, Rose Line Hofmann, Giuseppe Lonero,
Simona Muratori, Giuliana Papa, Francesca Santoni, Loretta Todesco,
Francesco Zanellati, Leili Khosravi

Rome, Italy.
E-mail: hm@interscalar.com

The paper introduces entropy analysis of bioelectrical activity based on harmonic signal distortion. Over a period of one year, the ultradian dynamics of the electrical activity of higher plants and the response on low frequency electromagnetic fields and modulated light were recorded. Biological active frequencies increasing or decreasing the entropy of bioelectrical activity were identified.

Introduction

Bioelectricity are electric potentials and currents produced by or occurring within living organisms. The experiments of Luigi Galvani and Alessandro Volta in the 18^{th} century proved the connection between electricity and muscle contraction in frogs and other animals. Today, electrical effects originating in active cells of the heart and the brain are commonly monitored and analyzed for diagnostic purposes.

In 1873, John Burdon-Sanderson [1] discovered bioelectrical activity in the leaf of the Venus flytrap due to stimulation. Recent studies evidence that intracellular electrical signals serve for information transmission in plant cells [2]. Electrical signals have been shown to accompany many processes in plant life, including respiration [3], water uptake and transport [4], leaf movement [5] and stress response [6]. Electrical signals also play an important role in physiological activities e.g. gas exchange, pollination, fertilization and gene expression [7].

Plant tissue is a good conductor of electricity, so that electrical resistivity is used for quantification of root structures and functioning. Studies of the spatiotemporal characteristics of the electrical network activity of the root apex evidence the existence of excitable traveling waves in plants [8], similar to those observed in non-nerve electrogenic tissues of animals. Electrical activity is mostly observed in the transition zone of the root apex, and points to a possible physiological role of synchronized electrical activity in this region.

Stefano Mancuso [9] has found rising evidence that the root apex is the key to the intelligence of higher plants. He argues that plants use the root system as a complex network instead of a single powerful brain. The plant-neurobiological paradigm of Mancuso assumes that plants have electrical activity similar to neurological ones. Recent research evidences that plants are endowed with feeling [10], complex social relations and can communicate with themselves and with animals, show behaviors similar to sleeping and playing.

Obviously, not only higher plants show intelligent behavior, but also unicellular organisms. For example, the plasmodium of the slime mould physarum polycephalum has the ability to find the minimum-length solution between two points in a labyrinth – a kind of tasks we used to think only animals could perform. Physarum polycephalum shows cognition without a brain, but also without neurons at all [11].

It is well known that the boundary frequencies of the electrical activity of the human brain are common to other mammals [12]. Furthermore, the frequencies of electrical brain activity and the natural frequencies of the electromagnetic activity of the Earth's atmosphere [13] are of the same range. This coincidence suggests that the frequencies of electrical brain activity could be of more fundamental concern and not limited to mammalian neurophysiology and, perhaps, higher plants, being embedded in the electromagnetic environment of the Earth, operate with the same frequencies of electrical activity.

Mammals including human have electrical brain activity [14] of the Theta type in the frequency range between 3 and 7 Hz, of Alpha type between 8 and 13 Hz and Beta type between 14 and 37 Hz. Below 3 Hz the brain activity is of the Delta type, and above 37 Hz the brain activity changes to Gamma. It is the physical separation of different states of brain activity that is essential for its stability. The violation of this separation can cause neurological disorders. In the case of human neurophysiology, Theta-Alpha or Alpha-Beta violation can cause speech and comprehension difficulties [15], depression and anxiety disorders [16].

Hence, the stability of the frequency boundaries separating Theta activity from Delta, and Beta activity from Alpha and Gamma is essential for neurophysiological health. The frequencies 3.0 Hz, 8.2 Hz, 13.5 Hz and 36.7 Hz define the boundaries. What is so special about these frequencies?

In [17] we have shown that the ratios of the boundary frequencies of the brain waves approximate Euler's number and its square root. Being attractors of transcendental numbers, they allow avoiding any resonance between the brain wave boundaries and thus stabilize the central nervous system. Indeed, the natural logarithms of the ratios of the boundary frequencies are close to integer and half values:

$$\ln\left(\frac{8.2}{3.0}\right) = 1.00 \qquad \ln\left(\frac{13.5}{8.2}\right) = 0.50 \qquad \ln\left(\frac{36.7}{13.5}\right) = 1.00$$

Furthermore, in [18] we have shown that these boundary frequencies approximate integer powers of Euler's number relative to the natural frequencies of the proton and the electron:

$$\ln\left(\frac{8.2\,\mathrm{Hz}}{\omega_e}\right) = -46 \qquad \ln\left(\frac{13.5\,\mathrm{Hz}}{\omega_p}\right) = -53$$

where $\omega_e = 7.76344 \cdot 10^{20}$ Hz and $\omega_p = 1.42549 \cdot 10^{24}$ Hz are the angular frequencies of the electron and the proton:

$$\omega_p = \frac{\mathrm{E}_p}{\hbar} \qquad \omega_e = \frac{\mathrm{E}_e}{\hbar}$$

where $\mathrm{E}_p = 938.272$ MeV and $\mathrm{E}_e = 0.511$ MeV are the rest energies of the proton and the electron [19], and $\hbar$ is the reduced Planck constant.

The fact that the brain wave boundary frequencies fit with integer powers of Euler's number relative to the natural frequencies of the proton and the electron indicates that *quantum* physical stability of the frequency boundaries is essential for brain activity.

Similar frequencies we find also in the Earth's electromagnetic spectrum, for example the Schumann resonances. Solar X-ray bursts can cause their variations [20]. In this case, the fundamental 7.8 Hz increases up to 8.2 Hz reaching exactly the stable Theta-Alpha boundary. The second Schumann mode 13.5 Hz coincides precisely with the Alpha-Beta boundary. It is remarkable that solar activity affects this mode much less or does not affect it at all because of its Euler stability. The third Schumann mode currently has a frequency of 20.3 Hz and must increase to 22.2 Hz in order to reach the next island of electron stability. By the way, such an increase is observed recently.

The coincidence of the boundary frequencies of brain activity with Schumann resonances demonstrates how precisely the electrical activity of biological systems is embedded in the electromagnetic activity of the Earth. Important to know that Euler's number and its roots make possible this embedding, because they are attractors of transcendental numbers and form islands of stability. They allow for exchanging information between systems of very different scales – the biophysical and the geophysical. Considering the universality of this embedding, it is very likely that it includes also the bioelectrical activity of plants.

In 1892, Otto Haake [21] showed that light can trigger the bioelectrical activity of plants. Changes in the light conditions may trigger variation in the potential of the guard cell membrane. In 1923, Alexander Gurwich discovered the phenomenon of mitogenetic radiation of biophotons – ultraweak biophysical photon emissions – detected in the UV-range of the spectrum [22]. He observed that these emissions can accelerate cell proliferation. In 1979, Vlail Kaznacheev [23] demonstrated experimentally that IR-A and UV-A biophotons are carriers of intercellular communication. In 1994, Fritz Popp [24, 25] discovered the regulatory significance of coherent biophotonic emissions and of non-random lognormal distributions of physiological parameters. Therefore, we recorded not only the ultradian dynamics of the electrical activity of the plants, but also their bioelectrical response on modulated red and infrared light.

Due to the potential use of bioelectrical phenomena for indicating the physiological condition of plants in agricultural fields, there have been several attempts to analyze these signals and extract their features using statistical and signal processing methods [26].

In his book 'What is Life?', Erwin Schrödinger stated that life feeds on negative entropy, or negentropy [27]. Biosystems are indeed fare from thermodynamic equilibrium, and the second law of classic thermodynamics does not apply to them. Within the thermodynamics of open systems developed by Ilya Prigogine [28], entropy can only be exchanged and, like energy, can neither be generated nor eliminated. From this point of view, Schrödinger's negentropy is a local decrease of entropy that appears as a consequence of entropy exchange of the biosystem with the environment. The ability of lower the own entropy through entropy exchange with the environment seems to be a universal criterion of vitality.

Therefore, for bioelectric signal processing, we applied entropy analysis based on harmonic signal distortion. Over a period of one year, we recorded the ultradian dynamics of the electrical activity in leafs of Orchidaceae phalaenopsis, Aloe vera, Ocimum basilicum and Panax ginseng, including the response on low frequency electromagnetic fields and modulated light.

Methods

Approaches to the study of electrical activities in plants include intracellular and extracellular measurements. The latter can detect the electrical signals produced by the tissue, and is applicable to the monitoring of an individual plant. The bioelectric resting potential across a cell membrane is typically about 50 millivolts. As electrical signals in plants are weak, they usually must be amplified and the recording device must have a high input impedance [29]. Therefore, for recording the bioelectrical signals in plant tissues we used a digital oscilloscope and attached the measuring electrode to a leaf.

For the purpose of shielding against uncontrolled external electromagnetic sources during the measurement, we placed the plant or the leaf in a container made of 1/16 aluminum sheet, similar to the described in [30] polyhedrons. Inside the container we placed also a coil generating a low frequency electromagnetic field. Figure 1 illustrates the experimental setup. Modulated red LED-light we applied as well. For light and field modulation, we chose the brain activity boundary frequencies and further frequencies of electron and proton stability in the range from 3 Hz to 15 kHz.

In [31] we have shown that destabilizing parametric resonance in oscillating systems of any complexity can be avoided

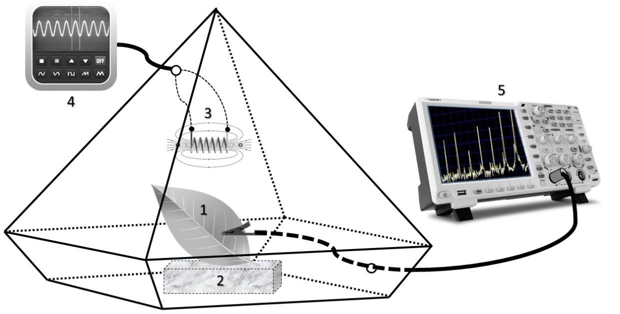


Fig. 1: The experimental setup: The plant or the leaf (1) was placed on a wooden platform (2) in a polyhedral container made of 1/16 aluminum sheet. Inside the container we placed also a coil (3) alimented by a frequency generator device (4). The bioelectrical signals were recorded by a digital oscilloscope (5).

if all frequency ratios correspond to integer powers of Euler's number. Essential for lasting stability in *real* systems is the prevention of proton and electron resonance. Therefore, also biosystems prefer frequencies corresponding to the electron or proton natural frequency divided by integer powers of Euler's number:

$$f_p = \frac{\omega_p}{e^n} \qquad f_e = \frac{\omega_e}{e^m}$$

where f_p and f_e are frequencies of proton respectively electron stability. The exponents n, m are integer.

As we already mentioned, the brain wave boundary frequencies are of electron and proton stability. However, it may be that some frequencies of bioelectrical processes in plants surpass the range of brain waves. Therefore, we applied also higher frequencies of electron and proton stability for field and light modulation (table 1).

n	$f_p = \omega_p/e^n$	m	$f_e = \omega_e/e^m$
46	15,011	39	8,965
47	5,522	40	3,298
48	2,032	41	1,213
49	747	42	446
50	275	43	164
51	101	44	60
52	37	45	22
53	14	46	8
54	5	47	3

Table 1: Frequencies (rounded) of proton f_p and electron f_e stability, which we applied for field and light modulation, and the corresponding integer exponents n, m of Euler's number.

For the purpose of control, we recorded the bioelectrical activity of the same plant or leaf alternately inside and outside the container. In a dark room, we applied also red and infrared light emitted by LEDs having 660 nm and 850 nm peak wavelengths, which was modulated by the same frequencies of electron and proton stability (table 1).

The measuring electrode of the oscilloscope picked up the bioelectrical signal directly from the leaf (fig. 1). The internal FFT-processor of the oscilloscope automatically stored the frequencies and amplitudes (voltages) of the harmonics to built-in memory. Based on the frequencies and amplitudes of the first 8 – 16 harmonics (depending on the field and light modulation frequency), the harmonic distortion HD of the bioelectrical signal was calculated:

$$\mathrm{HD} = \frac{\sqrt{(V_2^2 + V_3^2 + \cdots + V_n^2)/n}}{V_1}$$

where V_n is the n^{th} harmonic voltage and V_1 is the fundamental component. For example, a pure symmetrical triangle wave has HD of 12%, a square wave has 48%, and a sawtooth signal possesses 80%.

In this way, the distortion of a waveform relative to a pure sinewave can be measured by splitting the output wave into its constituent harmonics and noting the amplitude of each relative to the fundamental. The HD indicates the degree of order – disorder associated with the frequency spectrum of a signal. Therefore, we interpret the HD in terms of Shannon's information entropy [32].

Shannon's idea of information is that the value of a communicated message depends on the degree to which the content of the message is surprising. If an event is very probable, it is no surprise; hence the transmission of such a message carries very little information. From this point of view, HD is surprising, because it violates the expected $1/n^2$ decrease of the amplitudes of higher harmonics.

In order to process the HD-calculation automatically, we wrote a software that reads the FFT-datafile directly from the oscilloscope and stores the calculated HD values on SSD.

Results

We started recording the ultradian dynamics of the HD of bioelectrical signals in leafs of Orchidaceae phalaenopsis, Aloe vera, Ocimum basilicum and Panax ginseng in May 2020. To date we made a total of 1014 measurements of the bioelectrical response on low frequency electromagnetic fields and modulated light of these plants alternately inside and outside the shielding container (fig. 1).

The HD of the bioelectrical signals we measured varied between 67 and 88%. Figure 2 shows the ultradian dynamics of the HD measured on leafs in laboratory outside the container under conditions of natural illumination. The ultradian dynamics of HD measured on a leaf of the Orchidaceae phalaenopsis shows the typical increase in HD around noon and the decrease at sunset (fig. 2a) under otherwise constant environmental conditions. Fig. 2b shows clearly the continuous decrease of HD immediately after the weekly watering of Panax ginseng at 10 am. Even if it rained, but the plant in the laboratory did not get any water, the HD declined slightly. The increase of HD after 10 am in fig. 2c coincides with a powerful thunderstorm. All investigated plants showed similar reactions of HD on thunderstorm.

EXPONENT	FREQUENCY, Hz	HD Basil, %	HD Aloe, %
E46	15,011	83	77
P39	8,965	75	81
E47	5,522	81	76
P40	3,298	74	83
E48	2,032	79	69
P41	1,213	77	80
E49	747	76	71
P42	446	68	74
E50	275	85	81
P43	164	76	84
E51	101	79	67
P44	60	78	83
E52	37	83	81
P45	22	77	82
E53	14	84	75
P46	8	68	82
E54	5	78	69
P47	3	72	81

Table 2: Frequencies applied for field modulation inside the container (fig. 1) and the corresponding daily HD minima for Ocimum basilicum and Aloe vera. In accordance with tab. 1, P-exponents indicate frequencies of *proton* stability while E-exponents indicate frequencies of *electron* stability.

Fig. 2d illustrates how the HD dynamics of Orchidaceae phalaenopsis follows the weather conditions. The decrease in HD during the first 2 hours coincides with increasing cloudiness and the minimum HD with 1 hour of rain. As the cloudiness decreases after the rain, the HD will increase until the plant has been watered. Immediately afterwards the HD falls to the daily minimum. This reaction of the HD to weather conditions confirms that Orchidaceae phalaenopsis as well as Panax ginseng like a humid atmosphere but do not like intense sunlight. Fig. 2e shows the HD dynamics of Ocimum basilicum at the same day. In contrast to Orchidaceae phalaenopsis, increasing cloudiness provokes a significant increase of the HD in the electrical activity of O. basilicum. As the cloudiness decreases after the rain, the HD decreases as well, and the watering causes only a 1% fluctuation of HD. During and after the sunset the HD continuously increases. Obviously, Ocimum basilicum and Aloe vera are light-loving plants and show a significant decrease in HD with moderate sunlight and an increase in HD with a lack of light. All these conformities evidence the suitability of HD measurements for estimating trends in bioelectrical activity of plants.

In addition to these measurements, we studied the ultradian dynamics of HD on the same plants inside the shielding container (fig. 1), where we installed a coil for generating weak electromagnetic fields modulated by brain wave boundary frequencies and other frequencies of electron and proton stability (tab. 1). Inside the container, the plants did show very simple ultradian dynamics of HD with only one minimum and no usual reaction on weather conditions. Fig. 2f shows the HD minimum of Ocimum basilicum at noon. The frequency 453 Hz was applied for field modulation. The signal was sinus.

It is remarkable that the plants showed stable changes in the daily HD minimum as a function of the modulation frequency. Table 2 shows the frequencies of proton (P) and electron (E) stability applied for field modulation inside the container and the corresponding daily HD minima for Ocimum basilicum and Aloe vera. Apparently, O. basilicum prefers P-frequencies, and A. vera E-frequencies. The application of modulated light lead to similar results.

Conclusion

In this paper we introduced HD analysis of bioelectric signals as method of entropy variation measurement that could be applied as an efficient alternative agronomic tool at the service of producers for decision support and as tool of food quality control. Our study evidences that HD analysis of bioelectrical signals is a reliable method for evaluating the vitality of higher plants.

It is very likely that the HD of a bioelectrical signal is not just a measure of its entropy, but a way of bioelectrical intercellular communication. In this case, the relatively high HD values we measured could turn out to be an indicator of information density. Perhaps, the relative amplitude of each harmonic encodes some biologically significant information. This possibility and the ability of plants to communicate with other organisms could be the subject of further research.

Acknowledgements

The authors are grateful to the Community of Living Ethics for permanent support on all stages of the study.

Submitted on September 3, 2021

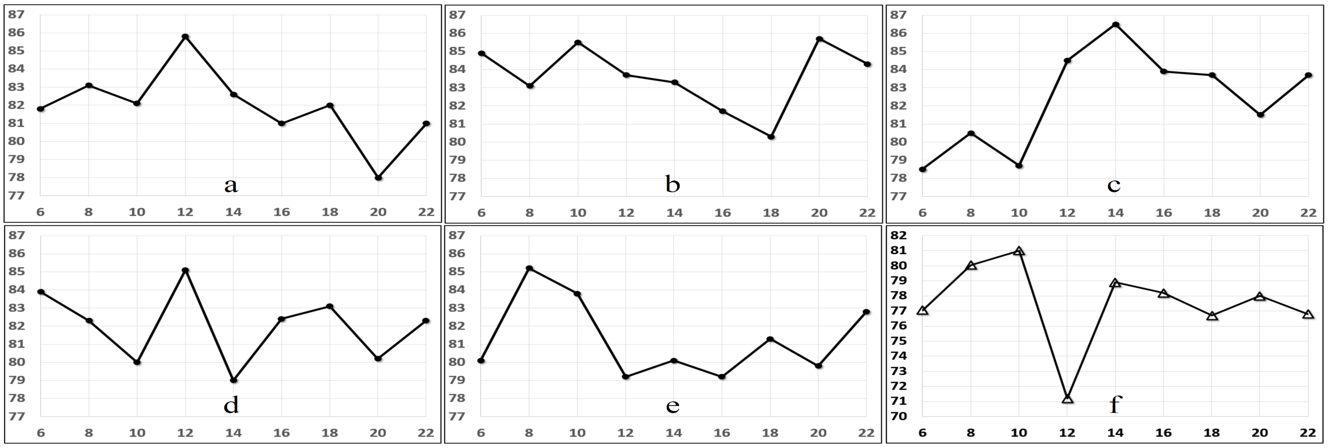


Fig. 2: Some examples of the ultradian dynamics of HD (vertical axes in %) measured on leafs in laboratory under conditions of natural illumination (a - e) and inside the aluminum container (f) in intervals of 2 hours starting at 6 am until 10 pm (horizontal axes). For detailed description, please read the main text.

References

1. Burdon-Sanderson J. Note on the electrical phenomena which accompany irritation of the leaf of Dionaea muscipula. *Proceedings of The Royal Society of London*, 1873, v. 21, 495.
2. Fromm J., Lautner S. Electrical signals and their physiological significance in plants. *Plant, Cell, Environment*, 2007, v. 30, 249–257.
3. Dziubinska H., Trebacz K., Zawadski T. The effect of excitation on the rate of respiration in the liverwort Conocephalum conicum. *Physiologia Plantarum*, 1989, v. 75, 417–423.
4. Davies E., Zawadski T., Witters D. Electrical activity and signal transmission in plants: how do plants know? In: Penel C., Greppin H. Plant signaling, plasma membrane and change of state. *University of Geneva*, 119–137, (1991).
5. Pickard B. G. Action potentials in higher plants. *The Botanical Review*, 1973, v. 39(2), 172–201.
6. Maffei M., Bossi S. Electrophysiology and plant responses to biotic stress. In: Volkov A. G. Plant electrophysiology – theory and methods. Springer-Verlag, 461–481, (2006).
7. Trebacz K., Dziubinska H., Krol E. Electrical signals in long-distance communication in plants. In: Baluska F., Mancuso S., Volkmann D. Communication in plants – neuronal aspects of plant life. Berlin, Heidelberg, *Springer-Verlag* p. 277–290, (2006).
8. Masi E. et al. Spatiotemporal dynamics of the electrical network activity in the root apex. *PNAS*, 2009, v. 106, no. 10, 4053.
9. Baluska F., Mancuso S., Volkmann D. Communication in Plants: Neuronal Aspects of Plant Life. ISBN 978-3-540-28516-8, (2006).
10. Calvo P., Sahi V. P., Trewavas A. Are plants sentient? *Plant, Cell, Environment*, 2017, v. 40, 2858–2869.
11. Nakagaki T. et al. Maze-solving by an amoeboid organism. *Nature*, 2000, v. 407, 470.
12. Stewart M., Fox S. E. Hippocampal theta activity in monkeys. *Brain Research*, 1991, v. 538(1), 59–63.
13. Constable C. Earth's Electromagnetic Environment. *Surveys in Geophysics*, 2016, v. 37, 27–45.
14. Tesche C. D., Karhu J. Theta oscillations index human hippocampal activation during a working memory task. *PNAS*, 2000, v. 97, no. 2, 919–924.
15. Kandana Arachchige Kendra G. et al. The Role of Iconic Gestures in Speech Comprehension: An Overview of Various Methodologies. *Frontiers in psychology*, 2021, v. 12, 634074.
16. Al-Ezzi A., Kamel N., Faye I., Gunaseli E. Review of EEG, ERP, and Brain Connectivity Estimators as Predictive Biomarkers of Social Anxiety Disorder. *Frontiers in Psychology*, 2020, v. 11, 730.
17. Müller H. Chain Systems of Harmonic Quantum Oscillators as a Fractal Model of Matter and Global Scaling in Biophysics. *Progress in Physics*, 2017, v. 13, 231–233.
18. Müller H. Global Scaling. The Fundamentals of Interscalar Cosmology. *New Heritage Publishers*, Brooklyn, New York, USA, ISBN 978-0-9981894-0-6, (2018).
19. Tanabashi M. et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018), www.pdg.lbl.gov
20. Roldugin V. C. et al. Schumann resonance frequency increase during solar X-ray bursts. *Journal of Geophysical Research*, 2014, v. 109, A01216.
21. Haake O. Über die Ursachen elektrischer Ströme in Pflanzen. *Flora*, 1892, v. 75, 455–487.
22. Gurvich A. G. Mitogenetic radiation. *Gosmedizdat*, Moscow, (1945).
23. Kaznacheev, V. P., Mikhailova, L. P., Radaeva, I. F. et al. Conditions for distant intercellular interaction during ultraviolet radiation. *Bulletin of Experimental Biology and Medicine*, 1979, v. 87, 492–495.
24. Popp F. A., Li K. H., Gu Q. Recent Advances in Biophoton Research and Its Applications. Singapore, (1992).
25. Chang J. J., Fisch J., Popp F. A. Biophotons. *Springer*, (1998).
26. Yan X. et al. Research progress on electrical signals in higher plants. *Progress in Natural Science*, 2009, v. 19, 531–541.
27. Schrödinger E. What is Life? *Cambridge University Press*, (1944).
28. Prigogine I., Nicolis G. Biological order, structure and instabilities. *Quarterly Reviews of Biophysics*, 1971, v. 4, 107–148.
29. Davies E. Electrical signals in plants. In: Volkov A. G. Plant electrophysiology. Berlin, *Springer-Verlag*, p. 407–422, (2006).
30. Müller H. et al. On the Acceleration of Free Fall inside Polyhedral Structures. *Progress in Physics*, 2018, v. 14, 220–225.
31. Müller H. On the Cosmological Significance of Euler's Number. *Progress in Physics*, 2019, v. 15, 17–21.
32. Shannon C. E. A Mathematical Theory of Communication. *Bell System Technical Journal*, 1948, v. 27(3), 379–423.

Ocean Currents and Tidal Movements: The Real Causes

Vladimir I. Danilov

Krasnoselsky, Kostroma, Russia. E-mail: danvlad@bk.ru

This article lists a pretty simple, radically different from other hypotheses, proved by measurement data, universal to any place at planet mechanism of causing currents and tides.

1 Introduction

Currents are classified on the basis of:

1. The causing forces (genetic classifications);
2. Stability;
3. Depth of location in the water column;
4. The nature of movement;
5. Physical and chemical properties.

There are three groups of currents:

- Gradient currents caused by horizontal gradients of hydrostatic pressure, arising when the isobaric surfaces are tilted relative to isopotential (level) surfaces
  - Density caused by horizontal density gradient
  - Compensatory caused by wind-induced sea level tilt
  - Barogradient caused by uneven atmospheric pressure over the sea surface
  - Seiche, resulting from seiche fluctuations in sea level
  - Stock, or wastewater, resulting from the emergence of excess water in any area of the sea (as a result of the influx of continental waters, precipitation, melting ice)
- Currents caused by wind
  - Drift caused only by the enticing action of the wind
  - Wind induced by both the pulling action of the wind and the inclination of sea level and changes in the density of the water caused by the wind
- Tidal currents caused by tides
  - Rebound current

All these types of currents do not explain the main permanent global current near the equator, directed from East to West, which, in the absence of continents, would be closed in round-like dust rings as in the atmosphere of Jupiter. (The nature of the equatorial countercurrent is not considered here. You can learn about it in the book *Equatorial countercurrents in the oceans* by V. B. Shtokman Leningrad 1948 [2].) .

2 A modern view of the causes of currents and tides

Sometimes the formation of currents is attributed to the Coriolis forces, while not taking into account that these forces are not real but conditional, used to describe different linear velocities of motion for points at different distances from the center on the radius, when the body rotates. In the case of the Earth's rotation, there is no movement of water along the radius, which can cause the appearance of such forces and such a constant movement of water around the circle.

Oceanic tides in modern scientific literature are considered as the rise of water due to the attraction from the Sun and the Moon, and at the same time they constantly try, using correction factors and various models, to lead to some kind of mathematics [1], considering that the Earth is, as it were, a body with its own vibration frequency. At the same time, forgetting that any oscillations have a decay time, and the processes under consideration last for many years. In fact, without identifying the main causes of the tides, this method is no better, and even worse, due to its complexity, a simple statistical table, that is, a method that has long been successfully used in the practice of navigation.

And the difference in the forces of gravity on an interval even of several kilometers (let's say that this is the depth of the ocean) at a distance of 380 000 km from the Moon, and 150 000 000 km from the Sun, cannot be so great as to cause the rise and movement of water. And this despite the fact that the entire mass of the Earth is nearby, which is much larger than the Moon.

The emphasis on tidal forces caused by the influence of the Sun and the Moon during the rotation of the Earth is made, for example, in the article [3], where a moving "hump" of the mantle allegedly causes the movement of water (discrete-wave motion). But it is not taken into account that the hump moves at a depth, and the main flows of the current do not fall below 200 meters [2], thus such a mechanism cannot work.

The action of tidal forces directly on the body of the ocean also cannot cause such a flow, for the reason that these forces act on masses of water, first from the East, and then in the same way from the West. Even if, which is not possible, they will first shift the mass of water in one direction, then they will return it back by the same amount.

Fig. 1 shows the current map and a conventional drawing [3] showing the similarity of the structures of large-scale currents of the Pacific, Atlantic and Indian oceans.

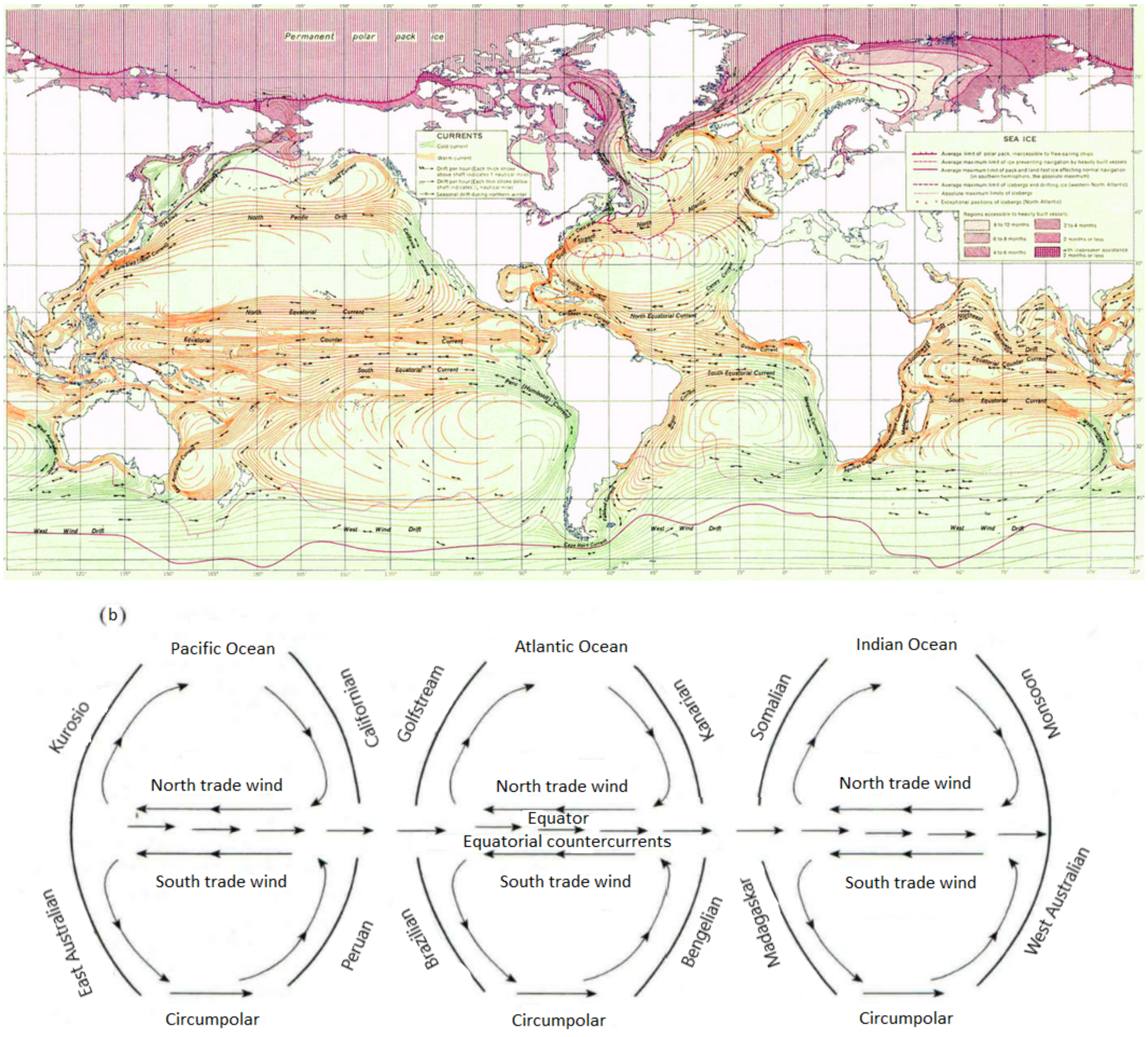


Fig. 1: Current map (1a) and a conventional drawing (1b) [3] showing the similarity of the structures of large-scale currents of the Pacific, Atlantic and Indian oceans.

3 The real reasons causing currents and tides

But there are currents and quite significant – the speed of movement is measured from 30 to 150 cm/s [2], which means there is a force that causes it. Moreover, this force is centuries old, constant direction. There are no external, observable forces. So there are internal ones.

Let's imagine the Earth as a kind of ball with a rather thin, relative to the total volume, shell, which can deform from the movement of the internal mass, if it is attracted to the external mass (Sun, Moon). Roughly it can be compared with an inflated air ball into which water has also been poured. Water, due to the force of gravity, will cause deformation of the shell, and when the ball rotates, this deformation will move in a circle. This is an analogue of the tide of a solid part of the Earth. But this is not an ocean tide! The high tide near the coast on the water will be caused by the ebb from the point of maximum rise of the mainland to the shores. If, for example, you pour water into a plastic plate and press from below, then the water will overflow to the edges. This fact is clearly visible when overlaying the graphs of the measured behavior of gravitational forces, the graph of the water level and the positions of the Sun and Moon at one measurement point.

In Fig. 2 (and also in Fig. 3), the maximum rise of water at high tides is clearly visible near the shores of the oceans. Fig. 2 shows measurement data at station "Posiet" of the Pacific coast, where

- On the horizontal axis Universal Time.
- Black is the measured force of gravity in μgal.
- Red is the Sun position.

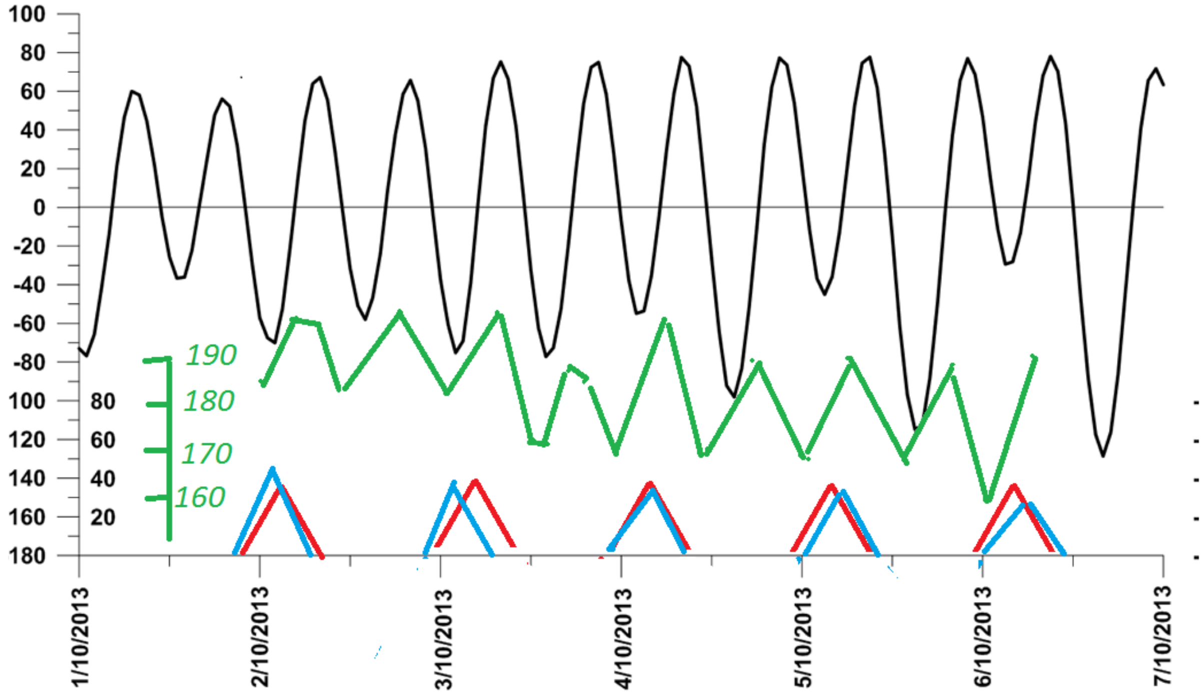


Fig. 2: Measurement data. Station "Posiet" of the Pacific coast.

- Blue is the position of the Moon in degrees above the horizon (time of sunrise, maximum position, sunset).
- Green is the ocean water level in cm.

The time interval is specially selected when the Sun and the Moon are close to the sky and at the same time affect the core of the Earth.

Measurement data of grav: the forces were provided by the staff of the laboratory of gravimetry of the POI FEB RAS. Station coordinates:

- ELLIPSOIDAL LATITUDE (DEGREE) 42.583
- ELLIPSOIDAL LONGITUDE (DEG. EAST) 131.158

Ocean level data: measurements taken at Posiet station, courtesy of the staff.

The data of the times of rise, maximum position, set and angle of ascent of the Sun and Moon were taken from the StarCalc program with reference to the station location.

It can be seen that a couple of hours before the passage of the Sun and the Moon to the zenith point, there is an ebb and flow of water and at the same time a decrease in the force of gravity, i.e. tide of the solid part of the planet. The ebb of water is also visible at night, when the tide of the mantle occurs from the departure of the planet's core to the opposite part of the Earth.

The time interval was specially selected when the Sun and the Moon are close and simultaneously affect the core of the Earth.

It is this fact that explains not the coincidence of the tides, but the coincidence of the ebb on the water with the positions of the Sun and the Moon at the zenith.

The "hump" on the mantle will change its position and size depending on:

- season (tilt of the axis of rotation);
- remoteness of the Moon and the Sun from the Earth;
- "dephasing", i.e. different positions between the Moon and the Sun;

then the tide near the coast will not be constant, but depend on these factors.

Now about the rise (tide) of the mantle on the opposite side of the globe. Unfortunately, it is difficult to demonstrate clearly, as in the first case, but even here everything is quite simple. The mass of the planet's core displaced towards the Sun and Moon will weaken the force of attraction on the opposite side of the ball in proportion to the square of the displacement distance. In the above graph, these will be the dips of the gravitational forces (black) during periods when there is neither the Sun nor the Moon above the measurement point. There is no other way to explain such a decrease in the forces of attraction, since the gravimeter reacts only to the force of attraction (mass).

In this way, multidirectional forces act on the gravimeter sensor:

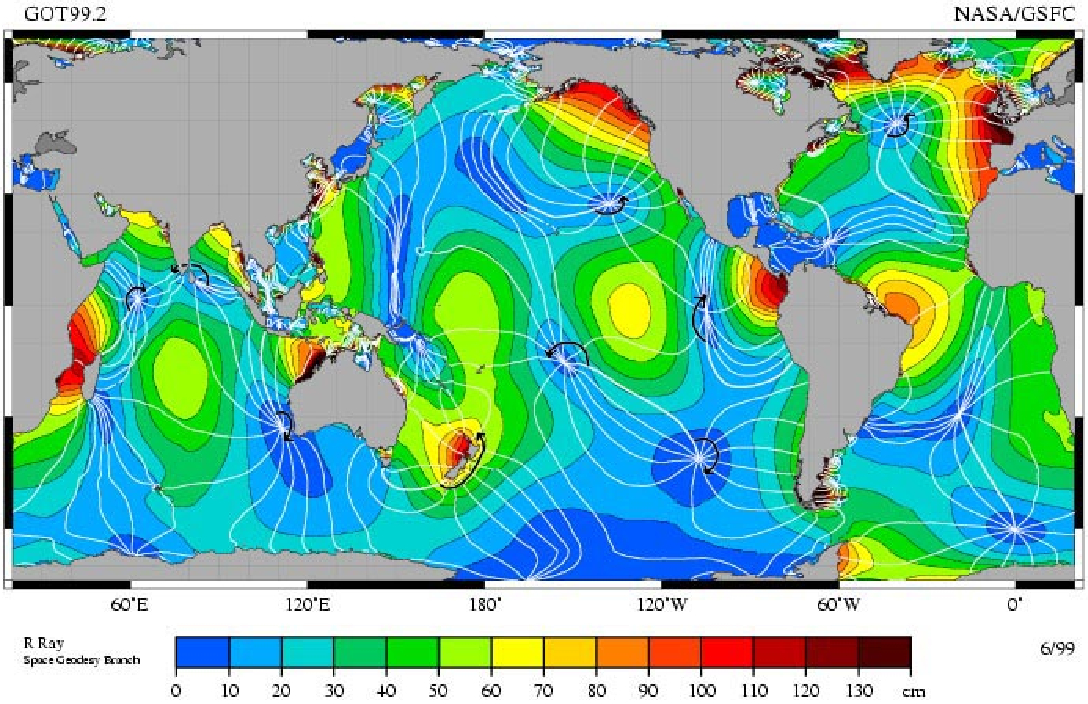


Fig. 3: Codial lines.

- attraction from the bulk of the Earth – constant force;
- attraction from the sun – variable force;
- attraction from the moon – variable force;
- attraction from the moving mass of the Earth's core – variable force, depending on the position of the Sun, Moon, season.

Based on the fact that there are many variables, mutually related quantities, calculating the mass of the moving part of the Earth's core is difficult (at least for me). When the Earth rotates, the "hump" will describe cyclic circular trajectories – this is the only observed movement in one direction, coinciding with the direction of movement of the main oceanic current [3]. The force of attraction of the mass of the moving inner core of the planet close to the water will force the mass of water to move in the same direction, but with a different force depending on the season. This is the reason for the main ocean current.

Since the bulk of the core moves in the equatorial region, the waters near the equator are also set in motion. Meeting the continents on its way, this current diverges to the sides of the equator and, since the basins of the oceans are practically closed, the water for the most part moves in a closed loop (see Fig. 1).

The change in the water level of the seas and oceans is only a demonstration of the change in the level of the solid surface of the planet. Water, due to the properties of fluidity, changes its level depending on the topography of the bottom and the coast. At the same time, the values of changes in the solid shell of the Earth depend on its structure and thickness. Mountain and continental massifs with large deep parts will naturally be less affected than lower, thinner, underwater areas. That is why the waters of the lakes practically do not change their level, since they are located on the body of massive continents and at the same time the level of the bottom of the entire reservoir changes slightly. On the plain of the oceans, amphidermal points (where there are no tides) and codial lines (lines connecting all points on the map where the crest of a tidal wave appears simultaneously, that is, points in which full water occurs simultaneously). If the tide would arise only from exposure to water, this could not be.

4 Conclusion

In addition to the processes under consideration, the movement of the planet's core leads to such consequences as the formation of the Earth's magnetic field, mountain building, continental drift, earthquakes, an astronomical shift relative to the reference time, *etc* [4].

Received on August 13, 2021

References

1. Melchior P. Earth Tides. MIR Publishing House, Moscow, 1968.
2. Shtokman V.B. Equatorial countercurrents in the oceans. Gidrometeoizdat, Leningrad, 1948.
3. Garetsky R.G., Dobrolyubov A.I. Tidal discrete-wave motions and continental drift. *Geotectonics*, 2016, v. 1, 3–13.
4. Danilov V,I. Consequences of the mutual influence of the planets. *Geophysical Bulletin*, 2017, v. 3, 11.

Searching for the 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

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. Although it's hard to believe, neither the original calculation nor the subsequent correction was ever published. Therefore, the entire prestige of QED and the Standard Model depend on the calculation of a single Feynman diagram (IIc) that has never been published and cannot be independently verified. In this article, we begin the search for any published recalculation of this Feynman diagram IIc that allows us to independently validate the theoretical calculation.

1 The big problem

1.1 Renormalization

The Standard Model of Particle Physics brings together two different physical theories: Electroweak Theory (EWT) and Quantum Chromodynamics (QCD). For decades, a "Grand Unification Theory (GUT)" has been unsuccessfully sought to integrate both theories into one unified theory.

Both QCD and EWT are mainly mathematical theories. The aim is to identify a set of gauge symmetries for each theory that allows a concrete mathematical formulation to be obtained. EWT forms a SU(2) × U(1) symmetry gauge group while QCD forms a SU(3) symmetry gauge group. The theory is considered correct if the theoretical values obtained with these mathematical formulas coincide with the experimental values obtained with particle colliders.

Both QCD and EWT are based on and completely dependent on the validity of quantum electrodynamics (QED), developed by Feynman, Schwinger, and Dyson. QED in turn is a quantum field theory (QFT). QFT emerged in the 1930s in an attempt to quantify the electromagnetic field itself. But QFT has a serious problem. All calculations give the same result: Infinity.

In the 1940s, QED developers managed to solve the infinities problem using a technique called "Renormalization". Many methods can be used to eliminate these infinities, but the main ones are:

- Substitution: replacing a divergent series with a specific finite value that has been arbitrarily chosen (for example, the energy of an electron).
- Separation: separating an infinite series into two components, one that diverges to infinity and another that converges to a finite value. Eventually, the infinite component is ignored and only the finite part remains.
- Cut-off: focusing on an arbitrary term in the evolution of a series that diverges to infinity and ignoring the rest of the terms of the series.

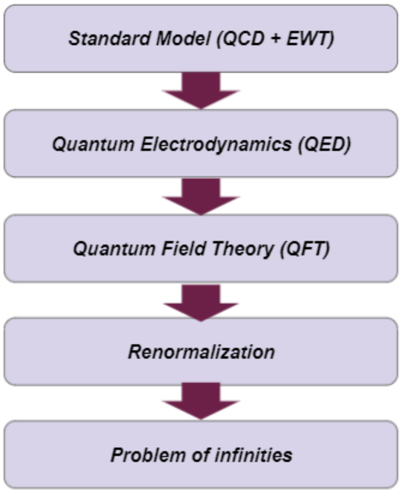


Fig. 1: Layers of logical dependencies.

As an example of the use of these Renormalization techniques we can look at the calculation of the Casimir effect [4]. The equation of the Casimir effect depends on the Riemann function. However, the Riemann function is defined only for positive values, since for negative values the Riemann function diverges to infinity.

$$\frac{F_c}{A} = \frac{d}{da}\frac{< E >}{A} = -\frac{\hbar c \pi^2}{2a^4}\,\zeta(-3) = \frac{\hbar c \pi^2}{20a^4}\,\zeta(-1) \qquad (1)$$

In particular the Riemann function of −1, $\zeta(-1)$, corresponds to the value of the sum of all positive integers. Applying a Renormalization technique, the Indian mathematician Ramanujan came to the conclusion that the sum of all positive integers is not infinity but $-1/12$ [3]. And this is precisely the value that is used in the equation of the Casimir effect.

$$\zeta(-1) = \sum_{n=1}^{\infty} n = 1 + 2 + 3 + 4 + 5 + ... = \frac{-1}{12} \qquad (2)$$

$$\frac{F_c}{A} = \frac{\hbar c \pi^2}{20a^4}\left(\frac{-1}{12}\right) = -\frac{\hbar c \pi^2}{240a^4} \qquad (3)$$

Despite being one of the main creators of QED, Feynman was not very convinced about Renormalization:

> The shell game that we play is technically called ‘renormalization’. But no matter how clever the word, it is still what I would call a dippy process! Having to resort to such hocus-pocus has prevented us from proving that the theory of quantum electrodynamics is mathematically self-consistent. It’s surprising that the theory still hasn’t been proved self-consistent one way or the other by now; I suspect that renormalization is not mathematically legitimate. [1]

For his part, Dirac was always clearly against these techniques:

> I must say that I am very dissatisfied with the situation because this so-called ‘good theory’ does involve neglecting infinities which appear in its equations, ignoring them in an arbitrary way. This is just not sensible mathematics. Sensible mathematics involves disregarding a quantity when it is small – not neglecting it just because it is infinitely great and you do not want it! [2]

Today, the scientific community accepts these renormalization techniques as fully legitimate. But if Dirac was right and renormalization is not a legitimate mathematical technique, then the Standard Model, EWT, QCD, QED and all theories based on QFT would be incorrect and worthless.

1.2 QED precision

The entire credibility of the renormalization techniques is based on its level of precision of the theoretical value with respect to the experimental value. As an example, the electron g-factor offers an impressive level of precision of 12 decimal places:

- Experimental value [12]: 1.001 159 652 180 73 (28),
- Theoretical value [13]: 1.001 159 652 182 032 (720).

In 1970, Brodsky and Drell summarised the situation in their paper *The present status of the Quantum Electrodynamics* as follows:

> The renormalization constants are infinite so that each calculation of a physical quantity has an infinity buried in it. Whether this infinity is a disease of the mathematical techniques of perturbation expansions, or whether it is symptomatic of the ills accompanying the idealization of a continuum theory, we don’t know. Perhaps there is a “fundamental length” at small distances that regularizes these divergences (...). Quantum electrodynamics has never been more successful in its confrontation with experiment than it is now. There is really no outstanding discrepancy despite our pursuing the limits of the theory to higher accuracy and smaller (...) however, and despite its phenomenal success, the fundamental problems of renormalization in local field theory and the nature of the exact solutions of quantum electrodynamics are still to be resolved. [14]

It seems inconceivable that using an incorrect theory, we can obtain the correct results with an unprecedented level of precision. And it is extremely unlikely that this finite theoretical value coincides with the experimental value by pure chance. Therefore, the only reasonable explanation is that renormalization techniques must be mathematically legitimate even though we cannot prove it at the moment.

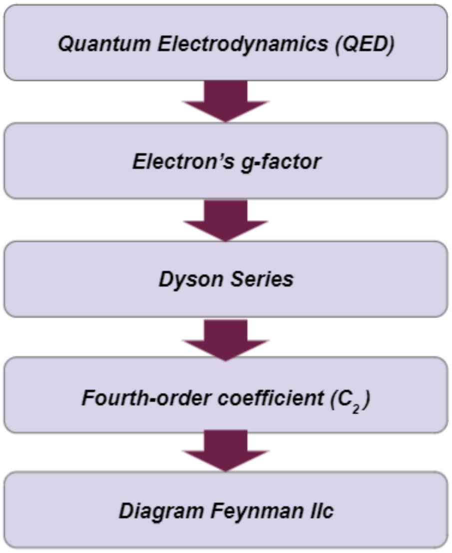


Fig. 2: Layers of logical dependencies.

1.3 Dyson series

Quantum Electrodynamics (QED) is considered the most accurate theory in the history of science. However, this impressive precision is limited to a single experimental value: the anomalous magnetic moment of the electron (g-factor).

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]. Each coefficient in the series requires the calculation of an increasing number of Feynman diagrams.

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

The first coefficient in the Dyson series is the Schwinger factor and has an exact value of 0.5. The second coefficient was initially calculated in 1950 by Karplus and Kroll [5], and it was corrected in 1957 by Petermann [6], who obtained a result of -0.328. The rest of the coefficients in the Dyson series were calculated many decades later with the help of supercomputers.

$$g = 1 + \frac{1}{2}\left(\frac{\alpha}{\pi}\right) - 0,328\left(\frac{\alpha}{\pi}\right)^2 = 1,0011596 \quad (5)$$

This result of the C_2 coefficient (fourth-order coefficient) of the Dyson series was decisive for the acceptance of the renormalization techniques proposed by Feynman, Schwinger, and Tomonaga, who received the Nobel Prize in 1965 for the development of QED. It can therefore be considered the most relevant theoretical calculation in modern physics.

1.4 Feynman diagram IIc

The error in the calculation of C_2 discovered by Petermann was found in the calculation of the Feynman diagram IIc.

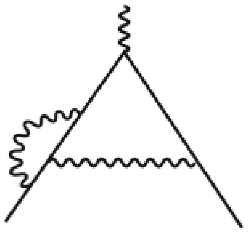


Fig. 3: Feynman diagram IIc.

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.

[Karplus & Kroll]

$$II_c = -\frac{323}{24} + \frac{31}{9}\pi^2 - \frac{49}{6}\pi^2 \ln(2) + \frac{107}{4}\zeta(3) = -3.178 \quad (6)$$

[Petermann]

$$II_c = -\frac{67}{24} + \frac{1}{18}\pi^2 + \frac{1}{3}\pi^2 \ln(2) - \frac{1}{2}\zeta(3) = -0.564 \quad (7)$$

However, hard to believe, neither the original calculation carried out in 1950 by Karplus and Kroll nor the subsequent correction of Petermann were ever published. Therefore, the entire legitimacy of the Standard Model and QED depends on the calculation of a single Feynman diagram (IIc) that has never been published and cannot be independently verified [11].

2 Searching for the missing calculation

2.1 Barbieri & Remiddi

At this point, we set out on a mission to find the missing calculation of the Feynman diagram IIc. We assume that given the seriousness of the situation, someone must have recalculated previously this Feynman diagram and published it years ago.

After a long search, we believe we found the paper we were looking for. It is a paper published in 1972 and written by Remiddi among other authors [8]. Remiddi is one of the most prestigious researchers in the calculation of the electron g-factor because in 1996 he published the definitive analytical value of the C_3 coefficient (sixth-order coefficient).

The paper is a long 93-page document entitled *Electron form factors up to fourth order*. It was published in 1972 by Barbieri and Remiddi. According to the authors:

> This paper is devoted to the analytic evaluation of the two form factors of the electromagnetic vertex of the electron in quantum electrodynamics, up to fourth order of the perturbative expansion (...) [Calculation] of the fourth-order form factors can also be found in the literature. They are the famous fourth-order anomalous magnetic moment evaluated by Petermann and Sommerfield (...). Such values are obviously reproduced in this paper. (...) Calculations are done in the framework of the usual Feynman-graph expansion of the S-matrix in the interaction representation, using the Feynman gauge for the photon propagator. The relevant graphs for second-order and fourth-order radiative corrections are shown (..). The approach is dispersive, and the discontinuities of the various Feynman graphs are obtained by means of the Cutkosky rules. [8]

From this introduction we understand that Barbieri and Remiddi performed a recalculation of the Feynman diagrams corresponding to the fourth-order coefficient (C_2) and they confirmed the results obtained by Petermann.

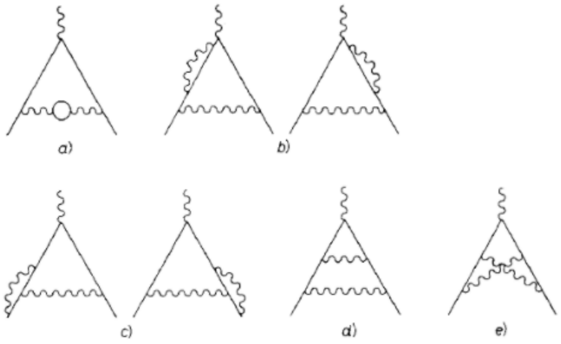


Fig. 4: C_2 Feynman diagrams.

The authors identify the Feynman IIc diagram as the "c" diagram in Fig. 4, divide it into two symmetrical diagrams and give it the descriptive name of "Corner Graphs". The result shown in the paper is identical to that published by Petermann in 1957.

In the 93 pages of the paper, the authors describe several of the techniques they have used to renormalize the divergences that appear in the calculations and how they have overcome the problems they have encountered. On the specific calculations, the authors state the following:

> Once these problems are mastered, a very long and complicated algebra is also needed to do in practice the calculation. Fortunately, the major part of it, like traces, straightforward algebraic manipulations, bookkeeping of analytic formulae, integrations by parts, differentiations and so on, was done by computer, us-

> ing the program SCHOONSCHIP of VELTMAN, without whose continuous and determinant help the present work could hardly have been accomplished. [8]

That is, they used a computer program to perform the mathematical calculations, but they did not publish the code of the programs used, so again, it is not possible to replicate the calculations.

Considering the date of the paper (1972) it is quite plausible to assume that there are no more calculations, since it was considered unnecessary to carry out more checks of the C_2 coefficient. Fortunately, in the paper itself, the authors identify two other independent calculations of the C_2 coefficient. One published in 1960 by Smrz and Uleha and the other published in 1962 by Terentiev.

2.2 Smrz & Uleha

We obtained the paper published in 1960 by these two Czech researchers [9]. It is a short paper of two pages where the situation generated in 1957 by Petermann's correction is explained. The paper indicates that the difference between the original Feynman IIc diagram calculation of Karplus and Kroll with respect to the one performed by Petermann is excessive. The authors state that they performed an independent calculation of the Feynman IIc diagram and obtained exactly the same result as Petermann.

> Since the considerable difference between the original value of the magnetic moment (Karplus & Kroll [5]) and the values calculated later (Petermann [6]) originates in the calculation of the contribution from the third diagram, only the value of this contribution was determined by the standard technique and the above regularization in the infra-red region. The contribution from the third diagram (-0.564) is in complete agreement with Petermann's value. [9]

Unfortunately, when searching for the reference of the work we note that it has not been published either: *Smrz P., Diploma thesis, Faculty of Tech. and Nucl. Physics, Prague 1960, unpublished.*

Just another unpublished paper claiming to have calculated the Feynman IIc diagram but with no one to review it.

2.3 Terentiev

We obtained a copy of the paper published by Terentiev in 1962. The paper contains about 50 pages [10]. The paper is only in Russian and there is no English translation. We identify the equation "60" of the paper as the C_2 coefficient of the Dyson series, with the same expression and value obtained by Petermann and Sommerfield.

$$g = 1 + \frac{1}{2}\left(\frac{\alpha}{\pi}\right) - 0,328\left(\frac{\alpha}{\pi}\right)^2 = 1,0011596 \tag{8}$$

Analyzing the document, we interpret that this equation is the result of the sum of nine other equations identified as equations 22, 24, 27, 31, 33, 47, 51, 58 and 59. There are nine equations instead of the five Feynman diagrams of Karplus and Kroll and none of the these equations correspond to the Feynman Diagram IIc.

However, it is not necessary to carry out a more in-depth analysis of the paper. On the first page of the Barbieri-Remiddi paper we can read a reference to Terentiev's paper:

> Actually, dispersion relations are used in the Terentiev work only to write down suitable multiple integral representations, which are in general manipulated to get the final result, without explicitly evaluating the discontinuities. The problem of infra-red divergences has been further overlooked, and many of the intermediate results are wrong, even if somewhat ad hoc compensations make the final result correct.

Year	Author	Status
1950	Karplus & Kroll	Wrong and unpublished
1957	Petermann	Right but unpublished
1957	Sommerfield	Right but using Green Functions instead of Feynman diagrams
1960	Smrz & Uleha	Right but unpublished
1962	Terentiev	Wrong intermediate results with ad hoc compensations to make the final result correct
1972	Remiddi	Right but unpublished computer calculation

Table 1: Fourth-order coefficient calculations.

3 Summary

Incredible as it may seem to believe, the most important calculation in the history of modern physics was published in 1950 by Karplus and Kroll and turned out to be completely incorrect. The error was not detected until seven years later by Petermann and Sommerfield. Neither the original calculation nor the subsequent correction was ever published. Therefore, the entire legitimacy of the Standard Model and QED depends on the calculation of a single Feynman diagram (IIc) that has never been published and cannot be independently verified.

In this paper we have detected three other published recalculations of the fourth-order coefficient of the g-factor. The detailed calculations of two of them were also not published (Barbieri-Remiddi and Smrz-Uleha). In the third calculation performed by Terentiev, serious errors were detected ten years after the original publication. Erroneous intermediate results manipulated with ad hoc compensations to obtain the correct final result.

Our search has been extensive, so we believe that there are no other published calculations of the Feynman IIc diagram. The only line of investigation that remains open would be to find the source code of the computer programs that are currently used to carry out this type of calculation.

Submitted on August 27, 2021

References

1. Feynman R. P. QED: The Strange Theory of Light and Matter. Princeton University Press, Pinceton, 1985. ISBN 0691024170.
2. Kragh H. Dirac: A scientific biography. Cambridge University Press, Cambridge, 1990. ISBN 0521017564.
3. Baez J. Euler's Proof That 1 + 2 + 3 + ··· = - 1/12. http://math.ucr.edu/home/baez/qg-winter2004/zeta.pdf, 2003.
4. Casimir H. B. G., Polder D. The Influence of Retardation on the London-van der Waals Forces. *Phys. Rev.*, 1948. v. 73 (4), 360–372.
5. 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.
6. Petermann A. Fourth order magnetic moment of the electron. *Helvetica Physica Acta*, 1957, v. 30, 407–408.
7. Sommerfield C. M. Magnetic Dipole Moment of the Electron. *Phys. Rev.*, 1957, v. 107 (1), 328–329.
8. Barbieri, Mignaco, Remiddi. Electron Form Factors up to Fourth Order. I & II. *Il Nuovo Cimento*, 1972, v. 1 (4), 824–916.
9. Smrz P., Ulehla I. Electrodynamic corrections to magnetic moment of electron. *Czech. Journ. Phys.*, 1960, v. 10, 966–968.
10. Terentiev M. V. *Zh. Eksp. Theor. Fiz.*, 1962, v. 43, 619.
11. Consa O. The unpublished Feyman Diagram IIc. *Progress in Physics*, 2020, v. 16 (2), 128–132.
12. 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.
13. 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.
14. Brodsky S. J., Drell S. D. The present status of the Quantum Electrodynamics. *Ann. Rev. Nucl. Part. Sci.*, 1970, v. 20, 147–194.

The Curse of Dimensionality in Physics

P. Reshma[1], P. Prasanth[2], and K. M. Udayanandan[3]

[1]Department of Physics, S N Polytechnic College, Kanhangad 671315, India. E-mail: reshmaperayil@gmail.com
[2]Department of Physics, Govt. Engineering College, Thrissur 680009, India. E-mail: prasanthpnairg@gmail.com
[3]Department of Physics, Associate Professor (Rtd), NAS College, Kanhangad 671314, India.
Corresponding author E-mail: udayanandan@gmail.com

The curse of dimensionality is a well-discussed issue in mathematics. Physicists also require n-dimensional space but because of the phase space choice, there is no need to worry about its consequences. This issue connected with dimensionality and related problems are discussed in this paper.

1 Introduction

We live in a 3-dimensional world and any dimension beyond this is called hyper-dimension. In the early decades of the 19th century, many articles were published listing works on "hyper-volume and surface" n-dimensional geometry. Swiss mathematician Ludwig Schlafli wrote a treatise on the subject in the early 1850's [1]. In 1858, a short description of this was translated into English by Arthur Cayley which gives the volume formula for an n-ball, commenting that it was determined long ago. In this paper, there were footnotes citing papers published in 1839 and 1841 [2] by the mathematician Eugene Catalan regarding descriptive geometry, number theory, *etc*. Though the earliest works encountered problems in computations, it was William Kingdon Clifford who published a solution in 1866 [3]. In the 1897 thesis, Heyl derived formulas for both volume and surface area and gives a clear idea of multidimensional geometry [4]. In 1911, Duncan Sommerville published a bibliography of non-Euclidean and n-dimensional geometry [5] giving the details on the works on hyper-sphere volumes. A book *An Introduction to the Geometry of N Dimensions* [6], by Duncan Sommerville, was published in 1929 which explains the n-ball formula and has a table of values for dimensions 1 to 7. In this paper, in the first section, we will give the formula for hyper-volume whose derivation is available in many statistical mechanics textbooks [7, 8]. In the other sections, we will discuss the so called "curse of dimensionality" and its consequences.

2 Hyper-volume

The n-dimensional volume of an Euclidean ball of radius R in n-dimensional Euclidean space [9] is

$$V_n(R) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} R^n \,, \tag{1}$$

where Γ is Euler's gamma function. The gamma function extends the factorial function to non-integer arguments. The volume of an n-dimensional sphere depends on the radius of the sphere (if we are considering the momentum space, the radius will be momentum) and the number of degrees of freedom. Now we want to know how the variation in n and R affects volume. For that, in the next section, we will numerically evaluate the variation of hyper-volume with increasing n for different radius.

3 The curse of dimensionality

We are all accustomed to live in low dimension spaces, mostly up to three dimensions. But relativity says we live in four dimensions [10] where the fourth dimension is time. String theory uses about ten dimensions [11, 12]. Our intuition about space can be misleading in high dimensions, rather more surprises awaits us there. Consider the case of an n dimensional sphere, and let us evaluate the volume for different dimensions for radius $R = 1$ and $R = 1.5$ which are given in Table 1. Initially an increase in volume is observed but later, volume decreases dramatically and almost approaches to zero at higher and higher dimensions. This effect is called the "curse of dimensionality" [13], often described as a phenomenon that arises when studying and using high-dimensional spaces. For $R = 1$, we can see that after reaching 5.26 the volume decreases, whereas for $R = 1.5$, after reaching 177.22, the volume decreases. These numbers depend on how the ratio $\pi^{\frac{n}{2}}/\Gamma(\frac{n}{2}+1)$ changes with n. Richard Bellman was the one who coined the term in 1957 [14, 15] when considering problems in dynamic programming.

In Fig. 1, we plot a graph with n along the x-axis and volume along the y-axis for ($R = 1$, 1.05, 1.10, 1.15, 1.20). In the graph, we can see that the volume first increases with n, reaches a maximum value for a particular value of n, called n_{max}. If we increase n further, the volume decreases. We can see that n_{max} shifts towards the right when R increases. All plots show that the volume of the n-ball vanishes to nothing as n approaches infinity.

4 What is really happening to the volume for large *n*?

First, we will check how the dimension will be influenced by the radius R. Taking the logarithm of the expression for the n-dimensional volume and applying Stirling's approximation in (1), we get

$$\ln V_n(R) \simeq \frac{n}{2}\,\ln\pi + n\ln R - \frac{n}{2}\ln\frac{n}{2} + \frac{n}{2}\,. \tag{2}$$

Dimension n	V_n for $R = 1$	V_n for $R = 1.5$
0	1	1
1	2	3
2	3.14	7.06
3	4.19	14.13
4	4.94	24.98
5	5.26	39.97
6	5.17	58.86
7	4.73	80.72
8	4.06	104.02
9	3.30	126.80
10	2.55	147.05
11	1.88	162, 97
12	1.33	173.24
13	0.91	177.22
14	0.59	174.94
15	0.38	167.03

Table 1: Values of hyper-volumes for $R = 1$ and $R = 1.5$.

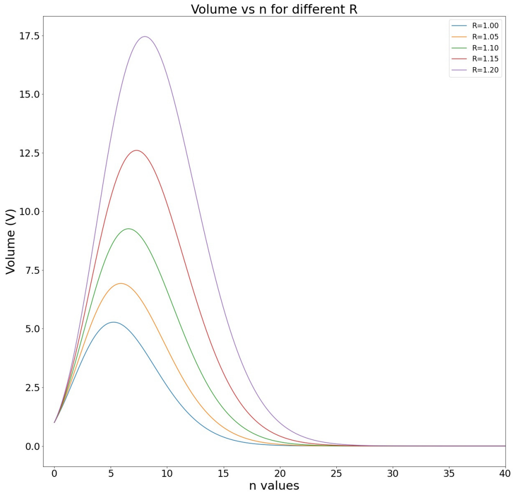


Fig. 1: A graph between n and volume for $R = 1, 1.05, 1.10, 1.15, 1.20$.

To find when the volume will decrease for different R, we take the derivative with respect to n of the above equation which gives

$$\frac{1}{V_n(R)}\frac{dV_n(R)}{dn} \simeq \frac{1}{2}\ln\pi + \ln R - \frac{1}{2}\ln\frac{n}{2}. \tag{3}$$

In order for the volume to be a maximum, $\frac{dV_n(R)}{dn}$ must be zero for a particular n. Hence we obtain

$$n_{max} \simeq 2\pi R^2 . \tag{4}$$

This relation of n_{max} for various R has a parabolic-type dependence which means the radius has no role in the decrease of volume. Next, we will find out what is happening to the volume for large n. There are arguments to show that data confined in the volume will be spreading to an outer shell for large n [16, 17]. Let us check whether this is true or not. For a sphere with radius ΔR less than R, the volume will be

$$V_n(R) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}(R-\Delta R)^n . \tag{5}$$

The volume of the shell will be given by subtracting (1)–(5). We calculated the volumes of n-dimensional spheres and shells for different n which is given in Table 2. A graph is also plotted with n along the x-axis and volumes of n-dimensional sphere and shell along the y-axis as in Fig. 2.

Dimension n	$V_n(R)$	$V_n(\Delta R)$
4	4.93	1.69
5	5.26	2.15
10	2.55	1.66
15	0.38	0.30
20	0.02	0.02
99	9.47×10^{-40}	9.47×10^{-40}
100	2.36×10^{-40}	2.36×10^{-40}

Table 2: Values of volumes of n-dimensional sphere and shell for different n.

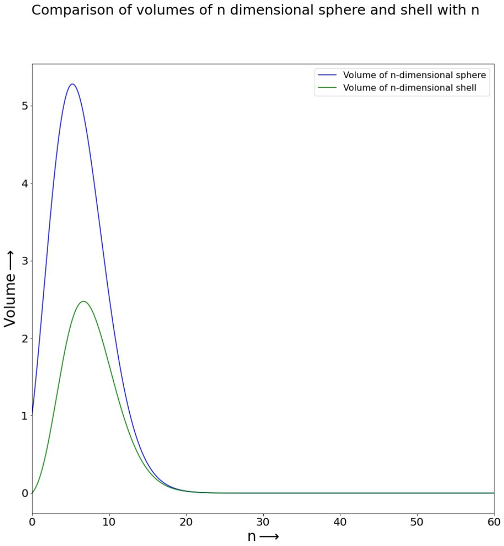


Fig. 2: A graph between n and volumes of n-dimensional sphere and shell.

Initially the volume of the shell is much less than the volume of the sphere. As n increases, both volumes decrease and become equal. We also found the percentage change in volume of the sphere to shell. The fraction of volume contained in the shell with thickness ΔR will be equal to

$$\text{Fractional volume} = \frac{\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}R^n - \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}(R-\Delta R)^n}{\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}R^n}\,.$$

On simplification,

$$\text{Fractional volume} = 1 - \left(1 - \frac{\Delta R}{R}\right)^n\,. \tag{6}$$

For $R = 1, \Delta R = 0.1$ and $R = 2, \Delta R = 0.1$, the fractional volumes in percentage are given in Table 3. All these show that the popular concept that the volume content is spreading into the surface area is not correct. Hence the curse of dimensionality remains unchanged.

Dimension	$R = 1$	$R = 2$
1	10	5
5	40.95	22.62
10	65.13	40.12
15	79.41	53.67
20	87.84	64.15
299	100	99.99
300	100	99.99

Table 3: Values of percentage of fractional volume for different n.

5 How Physicists overcome the curse using Statistical Mechanics

Statistical Mechanics (SM) provides the basis for many important branches of physics, including atomic and molecular physics, solid state physics, biophysics, astrophysics, environmental and socioeconomic physics. In statistical mechanics, we are interested in finding the thermodynamic properties of a system using n-dimensional space [18, 19]. It involves number of particles of the order of 10^{23} which are in continuous movement and hence have a continuous transformation in their position and momenta. So in order to predict the properties, we need to have information about all the possible values of position and momentum. For this, we construct a new space called a "phase space" which is a fusion of momentum and position spaces which is a six-dimensional space for N particles. In this space, the bridging equation to find the properties was given by Boltzmann [20, 21] as

$$S = k \ln \Omega \tag{7}$$

where S is the entropy, k is the Boltzmann constant and Ω is the number of available states in phase space which is given by [7, 22]

$$\Omega = \frac{V^{\frac{n}{3}}\pi^{\frac{n}{2}}}{h^n\,\Gamma\left(\frac{n}{2}+1\right)}R^n, \tag{8}$$

where V is the spatial volume and h is Planck's constant. Momentum volume of $3N$-dimension is [7, 22]

$$V_n = V_{3N} = \frac{\pi^{\frac{3N}{2}}R^{3N}}{\Gamma\left(\frac{3N}{2}+1\right)} = \frac{\pi^{\frac{n}{2}}R^n}{\Gamma\left(\frac{n}{2}+1\right)}\,. \tag{9}$$

In SM, we never have V_{3N} alone. We have both spatial volume V and momentum volume V_{3N} such that the total volume

$$V_{Total} = V^N V_{3N}\,. \tag{10}$$

But SM requires only Ω, the number of micro-states. Substituting for V_{Total}, we find the number of micro-states as

$$\Omega = \frac{V_{Total}}{h^{3N}N!} = \frac{V^N V_{3N}}{h^{3N}N!} = \frac{V^N\pi^{\frac{3N}{2}}R^{3N}}{h^{3N}N!\Gamma\left(\frac{3N}{2}+1\right)} \tag{11}$$

where $N!$ is used to avoid Gibbs paradox [7]. Simple calculations show that the number of micro-states (Ω) goes to infinity even for N just above 3 (Ω is of the order of 10^{1000} for $N = 100$). But because of the bridging equation, we require only $\ln\Omega$ and for that we carry out the following steps. Let us consider non-relativistic classical particles with energy $E = p^2/2m$. Then we have the radius of the momentum sphere $R = p = \sqrt{2mE}$ and we get

$$\Omega = \frac{\left(\frac{V}{h^3}\right)^N (2\pi mE)^{\frac{3N}{2}}}{N!\,\Gamma\left(\frac{3N}{2}+1\right)}\,. \tag{12}$$

Applying Stirling's approximation and carrying out suitable simplifications we arrive at

$$\ln\Omega \simeq N\ln\frac{V}{\lambda^3} - N\ln N + \frac{5}{2}N \tag{13}$$

where λ is the de Broglie wavelength. So we plot a graph between $\ln\Omega$ and N as in Fig. 3. The first graph shows a nonlinear variation because our choice of V/λ^3 is not realistic. In practice V/λ^3 will be always greater than 10^{25} and hence $\ln\Omega$ will be always proportional to N. This shows that in SM there is no need to worry about the decrease in volume of the n-dimensional space and we are not affected by the curse of dimensionality.

6 Conclusion

In statistical mechanics, in micro-canonical ensembles, we use the hyper-dimensional space to find the thermodynamic properties of a system. There are much literature [16, 17, 23, 24] showing that hyper-dimensional volume vanishes at large dimension or for large N. But this does not affect the properties of a system, which remains a paradox for physicists. This paradox is resolved in this paper. In SM, the classical particles are always in motion and hence to specify them we require both position and momentum simultaneously, which results in the phase space. We showed that because of the choice of the phase space, the curse of hyper-dimension is not affecting the properties and calculations in SM.

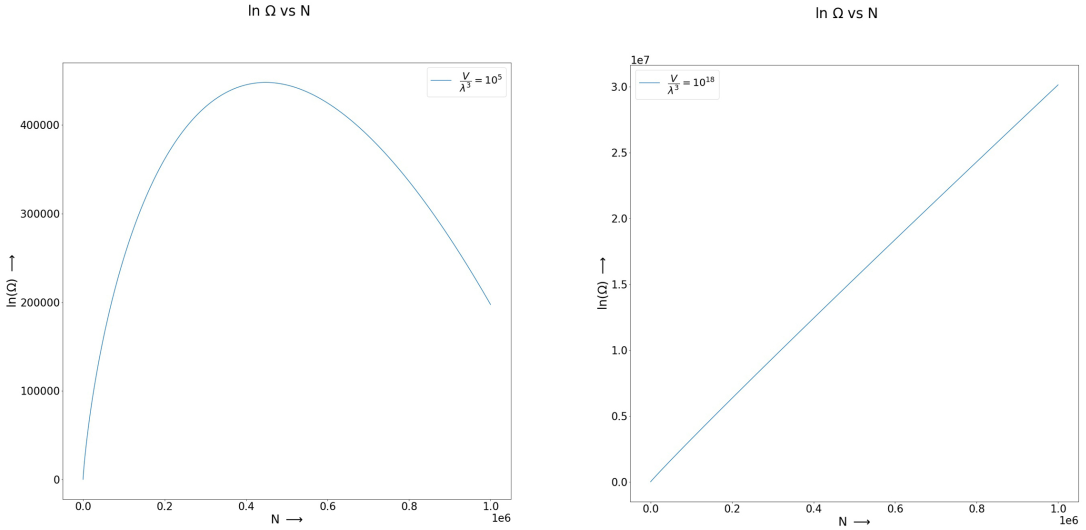


Fig. 3: A graph between $\ln\Omega$ and N for different V/λ^3.

Acknowledgements

During the preparation of the paper, the authors have been in constant discussions with Prof. K. P. N. Murthy, a great teacher of Physics from India, who died recently. As a token of respect to his inspiring comments, we wish to dedicate this paper to him. The authors also wish to thank Gautham Dathatreyan, Department of Science, Amrita Vishwa Vidyapeetham, Coimbatore for the computations using Python.

Received on August 18, 2021

References

1. Schläfli L. On a multiple integral. *The Quarterly Journal of Pure and Applied Mathematics*, 1858, v. 2, 269–301.
2. Catalan E. *Journal de Mathématiques Pures et Appliquées*, 1841, v. 6, 81–84.
3. Clifford W. K. *Mathematical Questions, with Their Solutions, from the "Educational Times"*, 1866, v. 6, 83–87.
4. Heyl P. R. Properties of the locus r = constant in space of n dimensions. *Mathematics*, 1897, v. 1, 33–39.
5. Sommerville D. M. Y. Bibliography of Non-Euclidean Geometry, Including the Theory of Parallels, the Foundation of Geometry, and Space of N Dimensions. Harrison and Sons, London, 1911.
6. Sommerville D. M. Y. An Introduction to the Geometry of N Dimensions. Dover Publications, New York, 1929.
7. Pathria R. K. and Beale P. D. Statistical Mechanics. Third Edition. Academic Press, Elsevier, 2011.
8. Kerson H. Introduction to Statistical Physics. Second edition. Chapman and Hall/CRC, 2009.
9. Callen H. B. Thermodynamics and An Introduction to Thermostatistics. 2nd ed. John Wiley and Sons, New York, 1985.
10. Florian C. Origins of Fourth Dimension Concepts. *The American Mathematical Monthly*, 1926, v. 33 (8), 397–406.
11. Jia B. Topological String Theory Revisited I: The Stage. *Int. J. Mod. Phys. A*, 2016, v. 31, 1650135.
12. Pincak R., Pigazzini A., Jafari S. and Ozel C. The "Emerging" Reality from "Hidden" Spaces. *Universe*, 2021, v. 7 (3), 75.
13. Donoho D. L. High-dimensional data analysis: The curses and blessings of dimensionality. In: Proceedings of the American Mathematical Society Conference on Math Challenges of the 21st Century. Los Angeles, CA, 2000.
14. Bellman R. Dynamic programming. Princeton University Press, Princeton, 1957.
15. Bellman R. E. Adaptive control processes: a guided tour. Princeton University Press, Princeton, 1961.
16. Ball K. An elementary introduction to modern convex geometry. In: Levy S., ed. Flavors of Geometry. Cambridge University Press, Cambridge, 1997.
17. Hayes B. An adventure in the Nth dimension. *American Scientist*, 2011, v. 99(6), 442.
18. Greiner I. W., Neise L., Stocker H. and Rischke D. Thermodynamics and Statistical Mechanics. Springer, 2001.
19. Bagchi B. Statistical Mechanics for Chemistry and Materials Science. Francis-Taylor, CRC Press, 2018.
20. Laird B. B. Entropy, Disorder, and Freezing. *J. Chem. Educ.*, 1999, v. 76, 1388–1390.
21. Schoepf D. C. A statistical development of entropy for the introductory physics course. *American Journal of Physics*, 2002, v. 70, 128.
22. Reif F. Fundamentals of Statistical and Thermal Physics. 1st ed. McGraw-Hill, Singapore, 1984.
23. Arfken G. B. Mathematical Methods for Physicists. Academic Press, Orlando, 1985.
24. Collins G. P. The Shapes of Space. *Sci. Amer.*, 2004, v. 291, 94–103.

Anomalous Magnetic Moment in Discrete Time

Young Joo Noh

E-mail: yjnoh777@gmail.com, Seongnam, Korea

The concept of causal delay in discrete time provides a correction for minimal coupling in electromagnetic interactions. This correction gives energy scale-dependent changes to the charge and mass of the elementary particles. As an application example of these results, this paper attempts to explain the anomaly of the g-factor. In particular, for the muon, $a = 0.001166$ is obtained from the approximation of its energy.

1 Introduction

In my last paper [1], I analyzed the effect of causal delay on the description of a dynamic system from a discrete time perspective. As a result, the dynamics system was divided into two worlds to which fundamentally different dynamical principles were applied, namely, type 1 and type 2. In the case of free particles, type 1 corresponds to ordinary matter that satisfies the existing relativistic quantum mechanics, and type 2 has properties similar to dark matter.

In this paper, I will discuss the interactions of type 1 particles, especially electromagnetic interactions. To do this, first, it is necessary to know the meaning of the result of type 1 in the case of free particles. In the case of free particles, the spinor state of type 1 satisfies the following equation

$$(x^\mu + \Delta x^\mu)\,\Psi(x) - x^\mu \Psi(x + \Delta x) = \Delta x^\mu \mathrm{e}^{-i\Delta x^\alpha P_\alpha}\Psi(x)\;. \tag{1}$$

The left side of (1) means the sum of contributions from $x - \Delta x$ and $x + \Delta x$ at x if the second term is translated by $-\Delta x$. The right-hand side means that the spinor change is only in phase, that is, the spinor forms a plane wave. Therefore, the above equation means that the sum of contributions from $x - \Delta x$ and $x + \Delta x$ forms a plane wave, that is, harmonic oscillation.

These facts require a new perspective on the matter field. In the existing field theory, especially quantum field theory, the field is based on the ontological basis of the statistical-mechanical analogy of the gathering of harmonic oscillators. From this point of view, harmonic oscillation is a property inherently immanent in the field. On the other hand, the matter field implied by (1) does not assume the property of harmonic oscillation inherent in nature. That is, harmonic oscillations are simply "formed" by the sum of contributions from the past and future of Δt_d. If we look at harmonic oscillations from this point of view, it can be said that interactions "deform" harmonic oscillations.

In the next section, I will discuss interactions in relativistic quantum mechanics with this new perspective on the matter field.

2 Modified Dirac equation

First, I will try to find the evolution operator equation for interacting particles corresponding to the evolution operator (1) for free particles. If the momentum at x and $x + \Delta x$ is p and $p+\Delta p$, and the spinor state is $\Psi(x, p)$ and $\Psi(x + \Delta x, p + \Delta p)$, respectively, the difference of cause-effect vector is as follows

$$\begin{aligned}
&(x^\mu + \Delta x^\mu)\,\Psi(x,p) - x^\mu \Psi(x+\Delta x, p+\Delta p)\\
&= (x^\mu + \Delta x^\mu)\,\Psi(x,p) -\\
&-x^\mu \sum_{m,n=0}^{\infty} \frac{1}{m!}\left(\Delta P^\alpha \frac{\partial}{\partial P^\alpha}\right)^m \frac{1}{n!}\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n \Psi(x,p)\\
&= (x^\mu + \Delta x^\mu)\,\Psi(x,p) -\\
&-\mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\, x^\mu \left(1 + \Delta x^\alpha \frac{\partial}{\partial x^\alpha} + \sum_{n=2}^{\infty} \frac{1}{n!}\left(\Delta x^\alpha \frac{\partial}{\partial x^\alpha}\right)^n\right)\Psi(x,p)\\
&= \left(\Delta x^\mu - x^\mu \Delta x \frac{\partial}{\partial x}\right)\Psi - \left(\mathrm{e}^{\Delta P \frac{\partial}{\partial P}} - 1\right)\left(x^\mu + x^\mu \Delta x \frac{\partial}{\partial x}\right)\Psi -\\
&-\mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\, x^\mu \sum_{n=2}^{\infty} \frac{1}{n!}\left(\Delta x \frac{\partial}{\partial x}\right)^n \Psi\\
&= \left\{\Delta x^\mu - \mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\left(x^\mu + x^\mu \Delta x \frac{\partial}{\partial x}\right) + x^\mu\right\}\Psi -\\
&-\mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\, x^\mu \sum_{n=2}^{\infty} \frac{1}{n!}\left(\Delta x \frac{\partial}{\partial x}\right)^n \Psi\\
&= \left\{x^\mu + \Delta x^\mu - \mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\,(x^\mu + \Delta x^\mu) - \mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\left[x^\mu, \Delta x \frac{\partial}{\partial x}\right]\right\}\Psi -\\
&-\mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\left[x^\mu, \sum_{n=2}^{\infty} \frac{1}{n!}\left(\Delta x \frac{\partial}{\partial x}\right)^n\right]\Psi\\
&= \left\{(x^\mu + \Delta x^\mu)\left(1 - \mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\right) - \mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\left[x^\mu, \mathrm{e}^{\Delta x \frac{\partial}{\partial x}}\right]\right\}\Psi(x,p)\\
&= \left\{(x^\mu + \Delta x^\mu)\left(1 - \mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\right) + \mathrm{e}^{\Delta P \frac{\partial}{\partial P}}\,\Delta x^\mu \mathrm{e}^{-i\Delta x \cdot P}\right\}\Psi(x,p)\;.
\end{aligned} \tag{2}$$

The right-hand side of (2) is the evolution operator of interacting particles. If we apply operator $\mathrm{e}^{-\Delta P \frac{\partial}{\partial P}}$ to both sides of (2) to get a simpler expression, it is as follows

$$\begin{aligned}
&(x^\mu + \Delta x^\mu)\,\Psi(x, p - \Delta p) - x^\mu \Psi(x + \Delta x, p)\\
&= \left\{\Delta x^\mu \mathrm{e}^{-i\Delta x \cdot P} + (x^\mu + \Delta x^\mu)\left(\mathrm{e}^{-\Delta P \frac{\partial}{\partial P}} - 1\right)\right\}\Psi(x,p)\;.
\end{aligned} \tag{3}$$

In (3), the first term on the right side is the evolution operator of a free particle and the second term is the interaction term. Since interaction is a local phenomenon, x in the interaction term can be set to Δx. And if

$$\mathrm{e}^{-\Delta \mathrm{P} \frac{\partial}{\partial \mathrm{P}}} \mathrm{e}^{\mathrm{i}\Delta \mathrm{x} \cdot \mathrm{P}} = \mathrm{e}^{-\mathrm{i}\Delta \mathrm{p} \cdot \Delta \mathrm{x}} \mathrm{e}^{\mathrm{i}\Delta \mathrm{x} \cdot \mathrm{P}} \tag{4}$$

is used, the final evolution operator O becomes

$$\mathrm{O}\Psi(x,p) = \left\{\mathrm{e}^{-i\Delta x \cdot P} + 2\left(\mathrm{e}^{-\mathrm{i}\Delta \mathrm{x} \cdot \Delta \mathrm{p}} - 1\right)\right\} \Psi(x,p) \,. \tag{5}$$

A little trick is needed here. Although not in practice, it is assumed that Ψ is analytic in order to maintain the conventional dynamics view. Then evolution operator O also needs to be defined as a locally continuous variable. What we need here is $\Delta x = x$. Therefore, the evolution operator O is equal to

$$\mathrm{O} = \mathrm{e}^{-ix \cdot P} + 2\left(\mathrm{e}^{-ix \cdot \Delta p} - 1\right) . \tag{6}$$

The evolution operator of free particles satisfies the Dirac equation. Therefore, the modified Dirac equation that the interaction evolution operator (6) must satisfy can be put as

$$\left(i\gamma^{\mu}\partial_{\mu} - m + A\right)\left\{\mathrm{e}^{-ix \cdot P} + 2\left(\mathrm{e}^{-ix \cdot \Delta p} - 1\right)\right\} \Psi(x,p) = 0 \,. \tag{7}$$

By finding A in (7), the following modified Dirac equation can be obtained

$$\mathrm{D_m}\Psi = \left(i\gamma^{\mu}\partial_{\mu} - f_{1r}\gamma^{\mu}p_{\mu} - f_{2r}\gamma^{\mu}\Delta p_{\mu}\right)\Psi = 0 \tag{8}$$

where

$$\begin{aligned} f_{1r} &= \mathrm{Re} f_1 = \frac{1}{3}\mathrm{Re}\frac{\mathrm{e}^{-ix \cdot p}}{\mathrm{e}^{-ix \cdot p} + 2\left(\mathrm{e}^{-ix \cdot \Delta p} - 1\right)} \\ f_{2r} &= \mathrm{Re} f_2 = \frac{1}{3}\mathrm{Re}\frac{2\mathrm{e}^{-ix \cdot \Delta p}}{\mathrm{e}^{-ix \cdot p} + 2\left(\mathrm{e}^{-ix \cdot \Delta p} - 1\right)} \,. \end{aligned} \tag{9}$$

In (9), the 1/3 factor is introduced under the condition that the sum of the coefficients in (8) is 0, that is $f_{1r} + f_{2r} = 1$. Now let's find the Hamiltonian using $\gamma^{\mu}p_{\mu} = m$. From (8),

$$\begin{aligned} \mathrm{H}\Psi &= i\partial_0\Psi \\ &= \left(-i\vec{\alpha} \cdot \vec{\nabla} + f_{1r}\beta m + f_{2r}\beta\gamma^{\mu}\Delta p_{\mu}\right)\Psi \\ &= \left\{\bar{\alpha} \cdot \vec{p} + f_{1r}\beta m + f_{2r}\left(\Delta p_0 - \vec{\alpha} \cdot \Delta\vec{p}\right)\right\}\Psi \\ &= \left\{\vec{\alpha} \cdot \left(\vec{p} - f_{2r}\Delta\vec{p}\right) + f_{1r}\beta m + f_{2r}\Delta p_0\right\}\Psi \,. \end{aligned} \tag{10}$$

Comparing the meaning of (10) with Hamiltonian $\mathrm{H}_0 = \vec{\alpha} \cdot \vec{p} + \beta m$ of free particles, it means that when there is a change in momentum and energy due to interactions, correction by $-f_{2r}\Delta\vec{p}$ and $-f_{2r}\Delta E$ is required, respectively. According to the existing minimal coupling theory, when a charge q interacts with an electromagnetic field, the momentum and energy become $\vec{p} - q\vec{A}$ and $E - q\phi$. Here, it can be inferred that the momentum change $\Delta\vec{p}$ and the energy change ΔE are $-q\vec{A}$ and $-q\phi$. Therefore, the combined momentum and energy of minimal coupling and causal delay are $\vec{p} - q\vec{A} + f_{2r}q\vec{A}$ and $E - q\phi + f_{2r}q\phi$. Rewriting, the resulting Hamiltonian is

$$\mathrm{H} - (1 - f_{2r})\,q\phi = \vec{\alpha} \cdot \left\{\vec{p} - (1 - f_{2r})\,q\vec{A}\right\} + \beta f_{1r}m \,. \tag{11}$$

Comparing (11) with the existing minimal coupling Hamiltonian, mass and charge can be newly defined as shown in (12) below. That is, the causal delay gives a modified mass and charge concept dependent on the energy scale:

$$\begin{aligned} m' &= f_{1r}m \\ q' &= (1 - f_{2r})\,q \,. \end{aligned} \tag{12}$$

3 Anomalous magnetic moment

The discovery of the Dirac equation made it possible to understand the property of spin of elementary particles, and predicted that the g-factor was 2. But, as a result of the measurement, anomaly exists, which was explained by a completely different paradigm of quantum field theory. However, according to the discussion in the previous section, considering the change in charge and mass due to the effect of causal delay, there is a possibility that the anomaly can be explained from the perspective of modified relativistic quantum mechanics.

When a particle with mass m' and charge q' is placed in an external field $A^{\mu} = \left(\phi, \vec{A}\right)$, the equation for calculating the magnetic moment is as follows

$$(\mathrm{H} - q'\phi)^2 = \left(\vec{p} - q'\vec{A}\right)^2 + m'^2 - q'\vec{\Sigma} \cdot \vec{B}. \tag{13}$$

where

$$\Sigma_j = \begin{pmatrix} \sigma_j & 0 \\ 0 & \sigma_j \end{pmatrix} .$$

If $q' = -e' = -(1 - f_{2r})\,e$, the nonrelativistic limit of (13) is obtained as follows

$$\mathrm{H} \cong m' + \frac{\left(\vec{p} + e'\vec{A}\right)^2}{2m'} - e'\phi + \frac{e'}{2m'}\vec{\Sigma} \cdot \vec{B} \tag{14}$$

where

$$\frac{e'}{2m'}\vec{\Sigma} \cdot \vec{B} = \frac{e}{2m}\frac{(1 - f_{2r})}{f_{1r}}2\vec{S} \cdot \vec{B} \equiv g\frac{e}{2m}\vec{S} \cdot \vec{B} \,. \tag{15}$$

Then, under the condition of $p \gg \Delta p$, the g-factor and the anomalous magnetic moment are as follows

$$\begin{aligned} \frac{g}{2} &= \frac{1 - f_{2r}}{f_{1r}} = 3 - 2\cos(x \cdot p) \\ a &= \frac{g}{2} - 1 = 2 - 2\cos(x \cdot p) \,. \end{aligned} \tag{16}$$

Previously local variable $x = \Delta x$. Then the phase value is

$$\Delta x \cdot p = E\,\Delta t - \vec{p} \cdot \Delta\vec{x} = \Delta t\left(E - \frac{\vec{p}^2}{m}\right). \tag{17}$$

In (17), Δt is the causal delay time. If $\Delta t = 0$, i.e. continuous time, there is no anomaly. Now let's define the physical meaning of Δt as follows.

Assumption: For a particle, the causal delay is the time it takes for light to pass through the particle's reduced Compton wavelength

$$\Delta t \equiv \frac{\lambda\!\!\!^{-}_c}{c} = \frac{\hbar}{mc^2}\,. \tag{18}$$

The Compton wavelength of a particle is a certain "region" of that particle. Therefore, the time it takes for light to pass through the region is the difference in time between cause and effect in the interaction between the particle and light.

Equation (18) is related to the "Penrose clock" [5]. According to him, *any individual stable massive particle plays a role as a virtually perfect clock.* And since $E = mc^2 = h\nu$, the frequency becomes $\nu = m\left(c^2/h\right)$. This can be said to be the same as (18). Since $\Delta t = 1/\nu$, the frequency becomes $\nu/2\pi = m\left(c^2/h\right)$. Therefore, it can be said that the causal delay time of massive particles plays the role of a fundamental clock that exists in nature.*

By the definition of Δt, the phase value of (17) can be calculated, and consequently the anomalous magnetic moment can be determined. An easy way to do this is the anomalous magnetic moment of the muon. In the case of the muon, at the so-called "magic momentum" $p_0 = 3.094$ GeV/c, the effect of the applied electric field for confinement of the muon is negligible. This means that the potential term in the energy value of (17) can be neglected. Therefore, the phase value of (17) is as follows

$$\begin{aligned}
&\frac{\vec{p}^2}{m} = \frac{E^2 - m_\mu^2}{m} = \frac{m_\mu^2\left(\gamma^2 - 1\right)}{m_\mu \gamma} = \beta^2 E \\
&\text{thus} \quad E - \frac{\vec{p}^2}{m} = \frac{E}{\gamma^2} = \frac{m_\mu}{\gamma} \\
&\text{and} \quad \therefore\ \Delta t\left(E - \frac{\vec{p}^2}{m}\right) = \frac{1}{\gamma}\,.
\end{aligned} \tag{19}$$

Therefore, the anomalous magnetic moment of the muon is

$$a_\mu = 2 - 2\cos\frac{1}{\gamma}\,. \tag{20}$$

Using (20) to find a_μ, γ corresponding to $E = 3.094$ GeV is 29.28, so

$$a_\mu = 0.001166\,. \tag{21}$$

Meanwhile, the value of a_μ recently announced by Fermilab is as follows.

$$a_\mu\,(\text{FNAL}) = 0.00116592040\,(54)\,. \tag{22}$$

The value of (21) is only calculated as an approximation of the muon energy. Thus, how accurately (20) predicts a_μ depends on the determination of γ, which is possible as an independent measurement of the cyclotron frequency $\omega_c = eB/m\gamma$.

4 Conclusions

Type 1 has a different view from the existing ones on the concept of field. It is that the harmonic oscillation of the field is formed by the sum of contributions from the past and future by Δt, not inherent in nature. And the interactions deform these harmonic oscillations. The result of analyzing the interaction from this point of view shows that it is more than the description of the interaction of the existing relativistic quantum mechanics.

The causal delay effect in discrete time corrects the existing minimal coupling theory, which leads to the result that the charge and mass of elementary particles change depending on the energy scale. As an example of such a result, it is partially shown that the anomalous magnetic moment of the muon can also be explained from this new perspective.

Received on September 8, 2021

References

1. Noh Y.J. Propagation of a Particle in Discrete Time. *Progress in Physics*, 2020, v. 16, 116–122.
2. Albahri T. *et al.* Measurement of the anomalous precession frequency of the muon in the Fermilab Muon g-2 Experiment. *Physical Review D*, 2021, v. 103, 072002.
3. Albahri T. *et al.* Magnetic-field measurement and analysis for the Muon g-2 Experiment at Fermilab. *Physical Review A*, 2021, v. 103, 042208.
4. Jegerlehner F. The Anomalous Magnetic Moment of the Muon. Springer, Heidelberg, 2007.
5. Penrose R. Cycles of Time. The Bodley Head, 2010.

---

*Except for (18) and this paragraph, the rest use natural units.

Representation of the Anomalous Magnetic Moment of the Muons via the Einstein-Podolsky-Rosen Completion of Quantum into Hadronic Mechanics

Ruggero Maria Santilli

The Institute for Basic Research, 35246 U. S. 19N, Suite 215, Palm Harbor, FL 34684, USA.
E-mail: research@i-b-r.org

In this paper, we briefly review half a century of research by various authors in the axiom-preserving completion of quantum mechanics into hadronic mechanics according to the 1935 Einstein-Podolsky-Rosen argument that "quantum mechanics is not a complete theory" (EPR argument). Said completion is intended to represent extended-particles in conditions of deep EPR entanglement with ensuing potential as well as contact non-Hamiltonian interactions represented by the new operator $\hat{T}$ in the associativity-preserving products $A \star B = A\hat{T}B$ of hadronic mechanics. We recall that muons are unstable and decay spontaneously with a mean live $\tau = 2.19703\times10^{-6}$ s, thus suggesting that they are composite, and therefore extended particles with constituents capable of triggering their decay. We then assume that the physical constituents of the muons are the ordinary electrons released in the spontaneous decay with low mode $\mu^- \rightarrow e^-, e^+, e^-$, resulting in the structure model according to hadronic mechanics (hm) $\mu^- = (\hat{e}^-_\downarrow, \hat{e}^+_\uparrow, \hat{e}^-_\downarrow)_{hm}$ where the "hat" characterizes iso-renormalizations due to non-Hamiltonian interactions. We show that the indicated hadronic structure model achieves an exact representation of *all* characteristics of muons, including rest energy, charge radius, mean life, spin, charge, spontaneous decays and anomalous magnetic moment.

1 Introduction

Recent, very accurate measurements [1] have established the following difference between the experimental value *muon* g-factor g_μ^{EXP} and its prediction via quantum electrodynamics g_μ^{QED}

$$\begin{aligned} & g_\mu^{EXP} - g_\mu^{QED} = \\ & = 2.00233184122 - 2.00233183620 \\ & = 0.00000000502 > 0\,. \end{aligned} \tag{1}$$

Additional accurate measurements [2] have shown deviations from quantum mechanical predictions for *atoms* in condensed matter, and measurements [3] have indicated bigger deviations from the predictions of quantum mechanics for *heavy ion.*

The above experiments support:

1) The validity of the historical 1935 argument by A. Einstein, B. Podolsky and N. Rosen that "quantum mechanics is not a complete theory" (EPR argument) [4];

2) The significance of historical completions of quantum mechanics, such as the non-linear completion by W. Heisenberg [5], the non-local completion by L. de Broglie and D. Bohm [6], and the completion via *hidden variables* by D. Bohm [7];

3) The validity of the recent verifications of the EPR argument by R. M. Santilli [8, 9] based on the completion of quantum mechanics (qm) into hadronic mechanics (hm) according to the EPR argument for the time-invariant representation of extended particles/wavepackets under potential as well as non-linear, non-local and non-potential interactions (see [10–12] for an outline of the basic methods, [13, 14] for recent overviews and [15–17] for detailed presentations).

2 Isotopic branch of hadronic mechanics

As it is well known, 20th century applied mathematics is characterized by a universal enveloping associative algebra ξ with conventional associative product $AB = A \times B$ between arbitrary quantities A, B, such as numbers, functions, operators, *etc*.

Lie algebras L with bracket between Hermitean operators $[A, B] = AB - BA$, then follow as the antisymmetric algebra attached to $L \approx \xi^-$, resulting in a unique characterization of Heisenberg's time evolution $idA/dt = [A, H]$ for point-particles under action-at-a-distance, potential interactions.

The EPR completion of quantum mechanics into hadronic mechanics has been studied to represent *extended* particles in conditions of mutual penetration, as occurring in the nuclear structure, with expected, additional, contact interactions of non-linear, non-local and non-potential type, hereon referred to as *non-Hamiltonian interactions.*

The axiom-preserving, thus isotopic branch of hadronic mechanics, known as *iso-mechanics,* and its underlying mathematics, known as *iso-mathematics,* represent the extended character of particles and their non-Hamiltonian interactions via the completion of the enveloping associative algebra ξ into the *universal enveloping iso-associative iso-algebra* $\hat{\xi}$ characterized by the *iso-product* (first introduced in Eq. (5), page 71 of [16] and treated in detail in [17])

$$A \star B = A\hat{T}B,\ \hat{T} > 0\,, \tag{2}$$

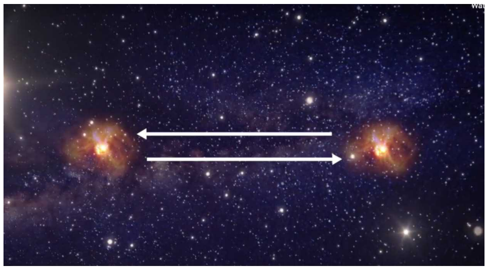


Fig. 1: In this figure, we illustrate the quantum entanglement of particles with instantaneous mutual actions at a distance, and recall the argument by A. Einstein, B. Podolsky and N. Rosen on the need for superluminal communications to represent said entanglement due to the local character of differentials, potentials and wavefunctions of the Schrödinger equation of quantum mechanics which solely allows a point-like characterization of particles [4].

where the quantity $\hat{T}$, called the *isotopic element,* is positive-definite but possesses otherwise an unrestricted dependence on all needed local variables (herein tacitly assumed).

The isotopic element has realizations of the type [12]

$$\begin{gathered}\hat{T} = \Pi_{\alpha=1,2}\,\mathrm{Diag}\left(\frac{1}{n^2_{1,\alpha}}, \frac{1}{n^2_{2,\alpha}}, \frac{1}{n^2_{3,\alpha}}, \frac{1}{n^2_{4,\alpha}}\right) e^{-\Gamma}, \\ n_{\mu,\alpha} > 0, \quad \Gamma > 0, \quad \mu = 1,2,3,4, \quad \alpha = 1,2,...,N,\end{gathered} \tag{3}$$

by charaterizing:

1) The dimension and shape of particles via semi-axes $n^2_{k,\alpha}, k = 1,2,3$ (with n_3 parallel to the spin) and the density $n^2_{4,\alpha}$, all n-characteristic quantities being normalized to the value $n^2_{\mu,\alpha} = 1$ for the vacuum.

2) Non-Hamiltonian interactions via the term e^{Γ}, where Γ is a positive-definite quantity with an unrestricted functional dependence on the wavefunctions as well as the characteristics of the medium in which the particles are immersed.

Iso-product (2) with realization (3) provides an explicit and concrete realization of Bohm's *hidden variables* [7], by therefore supporting the view that quantum mechanics does indeed admit hidden degrees of freedom, provided that quantum axions are realized in a way more general than that of the Copenhagen school.

It should also be noted that iso-product (2) and realizations (3) provide a quantitative representation of the completion of the conventional *quantum entanglement* of *point-like* particles under Hamiltonian interactions, into the covering *Einstein-Podolsky-Rosen entanglement* [13] which is applicable to *extended* particles with non-Hamiltonian interactions due to the deep overlapping of their wavepackets (see Figs. 1 and 2 and references quoted therein).

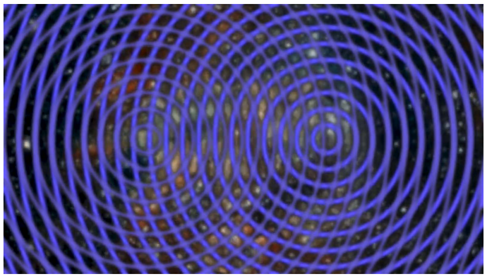


Fig. 2: In this figure, we illustrate the new Einstein-Podolsky-Rosen entanglement of particles introduced in [13], which is characterized by contact, therefore instantaneous and non-Hamiltonian interactions originating in the overlapping of the wavepackets of particles represented via isotopic elements of type (3). As such, the EPR entanglement prevents the applicability of Bell's inequality [18], allows an explicit and concrete realization of Bohm's hidden variables [7], and permits a progressive recovering of Einstein's determinism in the interior of hyperdense particles, with its full recovering at the limit of gravitational collapse [8, 9].

Despite its simplicity, iso-product (2) requires, for consistency, a compatible isotopy of the entire 20th century applied mathematics, with no exception known to the author [10]. In fact, iso-product (2) requires the following completions of 20th century applied mathematics (see Vol. I of [17] for a general treatment):

A) The compatible, completion of the basic unit $\hbar = 1$ of quantum mechanics into the integro-differential *iso-unit*

$$\hat{I} = 1/\hat{T} > 0, \quad \hat{I} \star A = A \star \hat{I} = A\,, \tag{4}$$

with ensuing completion of the conventional numeric field $F(n, \times, 1)$ of real $\mathcal{R}$, complex C and quaternionic Q numbers n into the *iso-fields* $\hat{F}(\hat{n}, \star, \hat{I})$ of *iso-real* $\hat{\mathcal{R}}$, *iso-complex* $\hat{C}$ and *iso-quaternionic* $\hat{Q}$ *iso-numbers* $\hat{n} = n\hat{I}$ and related isotopic operations [19] (see [20] for an independent study).

B) The completion of conventional functions $f(r)$ of a local variable r into *iso-functions* that, to have value on an iso-field, must have the structure [21] (see [22] for an independent study)

$$\hat{f}(\hat{r}) = [f(r\hat{I})]\,\hat{I}, \tag{5}$$

and related iso-differential iso-calculus [21] (see [23] for independent studies).

C) The completion of conventional spaces S over F into *iso-spaces* $\hat{S}$ over iso-fields $\hat{F}$ [21]. In particular, the conventional Minkowski space $M(x, \eta, I)$ over $\mathcal{R}$ with spacetime coordinates $x \in \mathcal{R}$, $x^4 = ct$, metric $\eta = \mathrm{Diag}(1,1,1,-1)$ and unit $I = \mathrm{Diag}(1,1,1,1)$, is mapped into the *iso-Minkowski iso-space* $\hat{M}(\hat{x}, \hat{\Omega}, \hat{I})$ over the iso-real iso-numbers $\hat{\mathcal{R}}$ [24] (see [25] for an independent study) with iso-coordinates $\hat{x} = x\hat{I} \in \hat{\mathcal{R}}$, iso-metric $\hat{\Omega} = (\hat{\eta})\hat{I} = (\hat{T}\eta)\hat{I}$, and iso-interval

$$\begin{gathered}(\hat{x}^\rho - \hat{y}^\rho)^{\hat{2}} = (\hat{x}^\rho - \hat{y}^\rho) \star \hat{\Omega}_{\rho\nu} \star (\hat{x}^\nu - \hat{y}^\nu) = \\ = \left[\frac{(x_1 - y_1)^2}{n_1^2} + \frac{(x_2 - y_2)^2}{n_2^2} + \frac{(x_3 - y_3)^2}{n_3^2} - \frac{(t_x - t_y)^2 c^2}{n_4^2}\right] \hat{I},\end{gathered} \tag{6}$$

where the exponential term $exp\{-\Gamma\}$ is imbedded into the n-characteristic quantities.

D) The compatible completion of all branches of Lie's theory first studied in [16] (see Vol. II of [17] for a general treatment and [26] for an independent study). For instance, an N-dimensional Lie algebra L with Hermitean generators X_k, $k = 1, 2, \ldots N$ is completed into the infinite family of Lie-Santilli iso-algebras $\hat{L}$ with iso-commutation rules

$$[X_i \hat{,} X_j] = X_i \star X_j - X_j \star X_i = C^k_{ij} X_k \,, \tag{7}$$

which iso-algebras are called *regular* or *irregular* depending on whether the structure quantities C^k_{ij} are constant or functions of local variables, respectively.

E) The completion of well-known space-time symmetries into the *iso-symmetries* of iso-space-time (6), including the completion of Lorentz's symmetry SO(3.1) into the *Lorentz-Santilli iso-symmetry* $\widehat{\mathrm{SO}}(3.1)$ [24] with iso-transformations

$$\begin{gathered} x^{1'} = x^1, \quad x^{2'} = x^2, \\ x^{3'} = \hat{\gamma}(x^3 - \hat{\beta}\tfrac{n_3}{n_4} x^4), \quad x^{4'} = \hat{\gamma}(x^4 - \hat{\beta}\tfrac{n_4}{n_3} x^3) \,, \end{gathered} \tag{8}$$

where

$$\hat{\beta}_k = \frac{v_k/n_k}{c_o/n_4}, \quad \hat{\gamma}_k = \frac{1}{\sqrt{1 - \hat{\beta}_k^2}} \,, \tag{9}$$

which provide the invariance of the local speed of light

$$C = \frac{c}{n_4} \,, \tag{10}$$

with consequential *iso-renormalization* of the energy (that is, renormalization caused by non-Hamiltonian interactions)

$$E = mc^2 \;\rightarrow\; \bar{E} = mC^2 = \frac{E}{n_4^2} \,. \tag{11}$$

Additionally, the isotopic representation of the anomalous magnetic moment of the muons requires the completion of the Lorentz-Poincaré symmetry P(3.1) into the *Lorentz-Poincaré-Santilli iso-symmetry* $\hat{\mathrm{P}}(3.1)$ [27], and the completion of the spinorial covering of the Lorentz-Poincaré symmetry $\mathcal{P}(3.1)$ into the *iso-spinorial covering of the Lorentz-Poincaré-Santilli iso-symmetry* $\hat{\mathcal{P}}(3.1)$ (in view of the spin 1/2 of the muons) [28] (see [11] for a recent review and [25] for independent studies).

We should also recall that all aspects of regular iso-mathematics and iso-mechanics can be constructed via the simple *non-unitary transform*

$$UU^\dagger = \hat{I} \neq I \,, \tag{12}$$

of *all* conventional mathematical or physical aspects, under which the unit of quantum mechanics is mapped into the iso-unit of iso-mechanics

$$\hbar = 1 \rightarrow U1U^\dagger = \hat{I} \,, \tag{13}$$

the conventional associative product AB is mapped into the iso-product

$$AB \rightarrow U(AB)U^\dagger = (UAU^\dagger)(UU^\dagger)^{-1}(UBU^\dagger) = \hat{A}\hat{T}\hat{B} \,, \tag{14}$$

and the same holds for the construction of all remaining regular iso-theories.

Finally, we recall that the isotopic element $\hat{T}$ represents physical characteristics of particles. Hence, the invariance of its numeric value is important for the consistency and experimental verification of any iso-theory. Such an invariance does indeed occur under the *infinite class of iso-equivalence* of isotopic methods which is given by the isotopic reformulation of non-unitary transforms called *iso-unitary iso-transforms*

$$UU^\dagger = \hat{I} \neq I, \quad U = \hat{U}\hat{T}^{1/2}, \quad \hat{U} \star \hat{U}^\dagger = \hat{U}^\dagger \star \hat{U} = \hat{I} \,, \tag{15}$$

under which we have the numeric invariance of the iso-unit [29]

$$\hat{I} \rightarrow \hat{U} \star \hat{I} \star \hat{U}^\dagger = \hat{I}' \equiv \hat{I} \,, \tag{16}$$

and of the isotopic element

$$\hat{A} \star \hat{B} \rightarrow \hat{U} \star (\hat{A} \star \hat{B}) \star \hat{U}^\dagger = \hat{A}'\hat{T}'\hat{B}', \quad \hat{T}' \equiv \hat{T} \,. \tag{17}$$

3 The structure of muons

As it is well known, the standard model assumes that muons $\mu^\pm$ are *elementary particles,* under which assumption, the sole known possibility of representing deviation (1) is the search for new particles and/or new interactions.

In this paper, we study the view presented on page 849 of the 1978 paper [30] (see also Section 2.5.5, page 163 of [12] for a recent update) according to which *muons are naturally unstable, and therefore they are composite, with a structure suitable to trigger their decay.*

Muons were then represented in [12] as a hadronic bound state of particles produced free in the spontaneous decays with the lowest mode $\mu^- \rightarrow e^-, e^\pm, e^-$ (tunnel effect of physical constituents), resulting in the three-body hadronic structure model with ordinary electrons

$$\mu^- = (\hat{e}^-_\downarrow, \hat{e}^+_\uparrow, \hat{e}^-_\downarrow)_{hm}, \tag{18}$$

in which the presence of positrons was instrumental for the representation of the muon spontaneous decays and its mean life.

Note that the constituents of model (18) are iso-electrons $\hat{e}^\pm$, rather than ordinary electrons $e^\pm$, due to their contact, non-Hamiltonian interactions due to their deep EPR entanglement (Fig. 3), which requires their characterization via an iso-irreducible iso-unitary iso-representation of $\hat{\mathcal{P}}(3.1)$ [28].

Note also that, since all constituents have point-like charges, the charge radius of the model is given by the radius of the orbit of the peripheral iso-electrons.

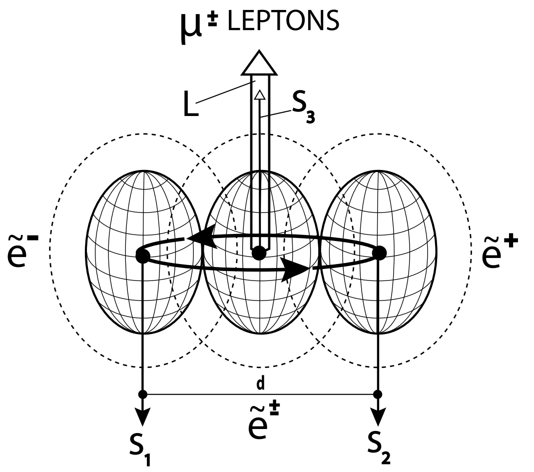


Fig. 3: In this figure, we illustrate structure model (18) of muons in its ground state with $L - 0$, first derived in the 1978 paper [30] (see page 849 on and Section 2.5.5, page 163 of the recent update [12]) as a three-body hadronic bound state of ordinary electrons produced free in the spontaneous decay with the lowest mode. The dashed lines represent the 1 fm wavepackets of the electrons and their overlap represents their deep EPR entanglement. Note the stability of model (18) due to the singlet couplings of the pairs of constituents, and the presence of a positron representing the muon spontaneous decay and its mean life via electron-positron annihilation.

The iso-Schrödinger equation of iso-mechanics, Eq. (5.2.4) of [30] (see also Eqs. (54), page 165, [12]) allowed the non-relativistic representation (here omitted for brevity) of the mass $E = 105.658$ MeV, charge radius of about $R = d/2 = 10^{-13}$ cm, mean life $\tau = 2.19703{\times}10^{-6}$ s, spin $S = 1/2$, charge $Q = \pm e$, and parity.

The 1978 paper [30] provided a representation of the magnetic moment of the muons known at that time, namely, that equal to the magnetic moment of the central electrons. This is due to the lack of contribution to the total magnetic moment from the electron-positron pair in the model which has a null charge and magnetic moment.

4 Anomalous magnetic moment of the muons

By keeping in mind that electrons have a point-like *charge* structure (but they have an extended wavepacket with radius of about 1 fm), it appears that the EPR entanglement of the constituents of model (18) (Fig. 3) causes a very small deformation (called *mutation* [30]) of the electrons such as to produce deviation (1). Its quantitative representation can be preliminarily achieved via the following isotopic procedure.

The relationship under isotopies between the magnetic moment and the spin of charged particles has been identified in Eq. (6.5), page 190 of [28], with ensuing relation for the g-factors (here reproduced for brevity without its derivation)

$$\hat{g}_\mu^{EXP} = \frac{n_4}{n_3} g_\mu^{QED} . \tag{19}$$

From value (1), we can then write

$$\frac{n_4}{n_3} = 1.00000000502 . \tag{20}$$

Model (18) for the structure of the muons has been indicated because it is necessary to identify the individual values of the characteristic quantities n_k^2, $k = 1, 2, 3$, representing the dimension and shape of the muons, and n_4^2, representing its density, with normalization $n_\mu^2 = 1$, $\mu = 1, 2, 3, 4$ for conventional electrons and positrons.

Under the assumption of model (18), the total rest energy of the constituents is $E_{cons} = 3E_e = 1.533$ MeV, while the muon rest energy is given by the familiar value $E_\mu = 105.7$ MeV. This implies the excess energy

$$\Delta E = 105.7\,\text{MeV} - 1.533\,\text{MeV} = 104.167\,\text{MeV} , \tag{21}$$

under which the Schrödinger equation no longer admits physically meaningful solutions [12, 30].

Isotopic methods provide a *mathematical* representation of excess energy (21) via iso-renormalization (11) with numeric value of the density

$$n_4^2 = \frac{3E_e}{E_\mu} = \frac{1.533}{105.7} = 0.0149 , \tag{22}$$

under which the consistency of the Schrödinger equation is restored at the isotopic level [12].

Excess energy (21) can be *physically* represented e.g. via the kinetic energy of the peripheral constituents. It should be indicated that missing energy (21) also occurs in the synthesis of the neutron from the hydrogen in the core of stars [31], as well as, more generally, in the synthesis of hadrons from lighter particles [12].

The use of normalization

$$n_1^2 + n_2^2 + n_3^2 = 1 , \tag{23}$$

then provides the desired first approximation of the charge distribution and shape of muons

$$n_1^2 = n_2^2 \approx 0.4926, \quad n_3^2 \approx 0.0149 . \tag{24}$$

The above data confirm the expected very prolate character of structure model (18) due to the point-like character of the constituents.

5 Concluding remarks

In the author's view, the most important notion emerging from the preceding study is that of the Einstein-Podolsky-Rosen entanglement representing the instantaneous and continuous communications between extended particles due to the overlapping of their wavepackets, with ensuing non-Hamiltonian interactions represented by (2) and (3), whose consistent treatment required the construction of iso-mathematics and iso-mechanics [17].

In fact, the EPR entanglement has the following important implications:

1. It prevents the applicability of Bell's inequality [18] due to the presence of non-Hamiltonian interactions first studied in [8];
2. It provides an explicit and concrete realization of Bohm's hidden variables [7] in terms of the isotopic element first achieved in [9]; and
3. It permits a preliminary, yet numerically exact and time invariant representation of all characteristics of muons, including their anomalous magnetic moment [1].

In closing, there seems to be grounds for a new physics, with expected corresponding advances in chemistry and biology, via the axiom-preserving completion of the Copenhagen simplest possible realization of quantum axioms into their broadest possible realization suggested by hadronic mechanics [13].

Received on September 14, 2021

References

1. Miller J. P., de Rafael E. and Roberts B. L. Muon (g-2): experiment and theory. *Rep. Prog. Phys.*, 2007, v. 70, 795–881.
2. Fadel M., Zibold T., Decamps B., Treutlein Ph. Spatial entanglement patterns and Einstein-Podolsky-Rosen steering in Bose-Einstein condensates. *Science*, 2018, v. 360, 409. www.santilli-foundation.org/Basel-paper.pdf.
3. Schukraft J. Heavy-ion physics with the ALICE experiment at the CERN Large Hadron Collider. *Trans. R. Soc.*, 2012, v. A370, 917–932. royalsocietypublishing.org/doi/10.1098/rsta.2011.0469.
4. Einstein A., Podolsky B., and Rosen N., Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 1935, v. 47, 777. www.eprdebates.org/docs/epr-argument.pdf.
5. Heisenberg W. *Nachr. Akad. Wiss. Gottingen*, 1953, v. IIa, 111. link.springer.com/chapter/10.1007/978-3-642-70079-8_23.
6. Stanford Encyclopedia of Philosophy. Bohmian (de Broglie-Bohm) Mechanics. 2021. plato.stanford.edu/entries/qm-bohm/.
7. Bohm D. A Suggested Interpretation of the Quantum Theory in Terms of "Hidden Variables". *Physical Review*, 1952, v. 85, 166. journals.aps.org/pr/abstract/10.1103/PhysRev.85.166.
8. Santilli R. M. Isorepresentation of the Lie-isotopic SU(2) Algebra with Application to Nuclear Physics and Local Realism. *Acta Applicandae Mathematicae*, 1998, v. 50, 177–190. www.santilli-foundation.org/docs/Santilli-27.pdf.
9. Santilli R. M. Studies on the classical determinism predicted by A. Einstein, B. Podolsky and N. Rosen. *Ratio Mathematica*, 2019, v. 37, 5–23. www.eprdebates.org/docs/epr-paper-ii.pdf.
10. Santilli R. M. Studies on A. Einstein, B. Podolsky, and N. Rosen prediction that quantum mechanics is not a complete theory. I: Basic methods. *Ratio Mathematica*, 2020, v. 38, 5–69. eprdebates.org/docs/epr-review-i.pdf.
11. Santilli R. M. Studies on A. Einstein, B. Podolsky, and N. Rosen prediction that quantum mechanics is not a complete theory. II: Apparent proof of the EPR argument. *Ratio Mathematica*, 2020, v. 38, 71–138. eprdebates.org/docs/epr-review-ii.pdf.
12. Santilli R. M. Studies on A. Einstein, B. Podolsky, and N. Rosen prediction that quantum mechanics is not a complete theory. III: Illustrative examples and applications. *Ratio Mathematica*, 2020, v. 38, 139–222. eprdebates.org/docs/epr-review-iii.pdf.
13. Santilli R. M. Overview of historical and recent verifications of the Einstein- Podolsky-Rosen argument and their applications to physics, chemistry and biology. APAV - Accademia Piceno Aprutina dei Velati, Pescara, Italy, in press, 2021. www.santilli-foundation.org/epr-overview-2021.pdf.
14. Dunning-Davies J. A Present Day Perspective on Einstein-Podolsky-Rosen and its Consequences. *Journal of Modern Physics*, 2021, v. 12, 887–936. www.scirp.org/journal/paperinformation.aspx?paperid=109219.
15. Santilli R. M. Foundation of Theoretical Mechanics. Vol. I, The Inverse Problem in Newtonian Mechanics. Springer-Verlag, Heidelberg, Germany, 1978. www.santilli-foundation.org/docs/Santilli-209.pdf.
16. Santilli R. M. Foundation of Theoretical Mechanics. Vol. II, Birkhoffian Generalization of Hamiltonian Mechanics. Springer-Verlag, Heidelberg, Germany, 1982. www.santilli-foundation.org/docs/santilli-69.pdf.
17. Santilli R. M. Elements of Hadronic Mechanics. Volumes I, II, III. Ukraine Academy of Sciences, Kiev, 1995 (I), 1995 (II), 2016 (III). www.i-b-r.org/Elements-Hadronic-Mechanics.htm.
18. Bell J. S. On the Einstein Podolsky Rosen paradox. *Physics*, 1964, v. 1, 195–200. cds.cern.ch/record/111654/files/vol1p195-200_001.pdf.
19. Santilli R. M. Isonumbers and Genonumbers of Dimensions 1, 2, 4, 8, their Isoduals and Pseudoduals, and "Hidden Numbers" of Dimension 3, 5, 6, 7. *Algebras, Groups and Geometries*, 1993, v. 10, 273–322. www.santilli-foundation.org/docs/Santilli-34.pdf.
20. Jiang C.-X. Foundations of Santilli Isonumber Theory. International Academic Press, Palm Harbor, FL, 2001. www.i-b-r.org/docs/jiang.pdf.
21. Santilli R. M. Nonlocal-Integral Isotopies of Differential Calculus, Mechanics and Geometries. In: Isotopies of Contemporary Mathematical Structures. *Rendiconti Circolo Matematico Palermo, Suppl.*, 1996, v. 42, 7–82. www.santilli-foundation.org/docs/Santilli-37.pdf.
22. Falcon Ganfornina R. M. and Valdes J. Fundamentos de la Isoteoria de Lie-Santilli. International Academic Press, Palm Harbor, FL, 2001. www.i-b-r.org/docs/spanish.pdf.
23. Georgiev S. Foundation of the IsoDifferential Calculus. Volumes I to V. Nova Academic Publishers, New York, NY, 2014 (I), 2014 (II), 2015 (III), 2015 (IV), (2016 (V).
24. Santilli R. M. Lie-isotopic Lifting of Special Relativity for Extended Deformable Particles. *Lettere Nuovo Cimento*, 1983, v. 37, 545–551. www.santilli-foundation.org/docs/Santilli-50.pdf.
25. Aringazin A. K., Jannussis A., Lopez F., Nishioka M. and Vel-janosky B. Santilli's Lie-Isotopic Generalization of Galilei and Einstein Relativities. Kostakaris Publishers, Athens, Greece, 1991. www.santilli-foundation.org/docs/Santilli-108.pdf.
26. Sourlas D. S. and Tsagas G. T. Mathematical Foundation of the Lie-Santilli Theory. Ukraine Academy of Sciences, Kiev, Ukraine, 1993. www.santilli-foundation.org/docs/santilli-70.pdf.
27. Santilli R. M. Nonlinear, Nonlocal and Noncanonical Isotopies of the Poincaré Symmetry. *Moscow Phys. Soc.*, 1993, v. 3, 255. www.santilli-foundation.org/docs/Santilli-40.pdf.
28. Santilli R. M. Recent theoretical and experimental evidence on the synthesis of the neutron. *Chinese J. System Eng. and Electr.*, 1995, v. 6, 177–186. www.santilli-foundation.org/docs/Santilli-18.pdf.

29. Santilli R. M. Invariant Lie-isotopic and Lie-admissible formulation of quantum deformations. *Found. Phys.*, 1997, v. 27, 1159–1177. www.santilli-foundation.org/docs/Santilli-06.pdf.

30. Santilli R. M. Need of subjecting to an experimental verification the validity within a hadron of Einstein special relativity and Pauli exclusion principle. *Hadronic J.*, 1978, v. 1, 574–901. www.santilli-foundation.org/docs/santilli-73.pdf.

31. Norman R., Beghella Bartoli S., Buckley B., Dunning-Davies J., J. Rak J., Santilli R. M. Experimental Confirmation of the Synthesis of Neutrons and Neutroids from a Hydrogen Gas. *American Journal of Modern Physics*, 2017, v. 6, 85–104. www.santilli-foundation.org/docs/confirmation-neutron-synthesis-2017.pdf.

LETTERS TO PROGRESS IN PHYSICS

On Eddington's Temperature of Interstellar Space and the Cosmic Microwave Background Radiation

Pierre A. Millette

E-mail: pierre.millette@uottawa.ca, Ottawa, Canada

We point out that there were several non-cosmological estimates of the blackbody temperature of interstellar space that predated and that were more accurate than the Cosmic Microwave Background (CMB) Big Bang estimates. They are disregarded and considered coincidental as they are not based on the cosmological Big Bang model. We note the importance of this question, as the energy requirements of the two different explanations (galactic vs cosmological) are substantially different. We also point out that the actual correct explanation can't be determined from the measurements done in our local neighbourhood inside the Milky Way.

> *The great tragedy of Science – the slaying of a beautiful hypothesis by an ugly fact.* Thomas Henry Huxley (1825–1895)*.

1 Introduction

Penzias and Wilson [1], while working at Bell Labs, measured an isotropic Microwave Background Radiation (MBR) of approximately 3 K, while using a sensitive antenna/receiver system under development. Initially, they thought the radio noise resulted from their equipment, but eventually they concluded that the background radiation was real.

Physicist Robert Dicke suggested that the background radiation was the Cosmic Microwave Background (CMB) radiation, believed to result from the Big Bang cosmological model. This interpretation was published in side-by-side letters by Penzias and Dicke in Astrophysical Journal Letters [2]. Penzias and Wilson measured an isotropic Microwave Background Radiation which became the Cosmic Microwave Background (CMB) radiation in the serendipitous communication with Dicke, nowadays the only accepted explanation for the measurement.

However, there were other earlier blackbody† temperature predictions, that were much closer to the initial measurement of Penzias and Wilson, than those from the Big Bang, but they were simply ignored as they did not originate from the Big Bang cosmological model. Interestingly enough, the very fact that the remarkably close blackbody temperature predictions do not originate in the Big Bang model is used against the validity of the other models in predicting a blackbody temperature in agreement with the Penzias and Wilson measurement!

At stake is whether the Microwave Background Radiation is universal and cosmic (i.e. CMB) or galactic in nature (i.e. MBR), with possibly every galaxy having slightly different local blackbody temperatures. The energy requirements of the two different explanations are substantially different. The reality is that this can't be determined from the measurements done in our local neighbourhood (at about 27 000 light-years from the galactic centre) within our Milky Way which is about 100 000 light-years across and about 2 000 light-years thick at the thin stellar disk that we are located in.

2 Eddington's "Temperature of interstellar space"

Assis and Neves in their 1995 paper *History of the 2.7 K Temperature Prior to Penzias and Wilson* [3] provide a review of earlier blackbody temperature determinations, prior to the Big Bang cosmological model temperature estimates of the late 1940s, 1950s and early 1960s which varied between 5 K and 50 K. Their conclusion that "the models based on a Universe in dynamical equilibrium without expansion predicted the 2.7 K temperature prior to and better than models based on the Big Bang" is, understandably so, not very popular.

The best-known earlier blackbody temperature prediction is that of $T = 3.2\,\mathrm{K}$ proposed by Arthur Stanley Eddington in 1926 [5], known as the *temperature of interstellar space* to clearly communicate that it is not related to the CMB, especially since Eddington's estimate was derived before the development of the Big Bang cosmological model. Modern commentators constantly remind us that it is coincidental and that it does not derive from the Big Bang model. We don't want people to see it as an explanation of the MBR that would be an alternative to the CMB Big Bang explanation!

Eddington, in his 1926 book *The Internal Constitution of the Stars* [6], further covered the topic in Chapter XIII, *Diffuse Matter in Space*. He computes an effective blackbody temperature of 3.18 K, but again, this has nothing to do with the 2.725 K blackbody spectrum of the Microwave Background Radiation (MBR), which we know is the Cosmic Microwave Background (CMB). Eddington states:

> The total light received by us from the stars is estimated to be equivalent to about 1000 stars of the

---

[*] Wikiquote. Thomas Henry Huxley. In his Presidential Address at the British Association in 1870, last modified 07:40 4 May 2019.

[†] The estimates are described as blackbody temperatures as the Stefan-Boltzmann blackbody radiation law was used to determine the temperature.

> first magnitude. ... We shall first calculate the energy density of this radiation. ... Accordingly the total radiation of the stars has an energy-density ... $E = 7.67 \cdot 10^{-13}$ erg/cm^3. By the formula $E = aT^4$ the effective temperature corresponding to this density is 3.18° absolute. [6, p. 371]

Eddington thus uses the Stefan-Boltzmann blackbody radiation law to determine the temperature of the blackbody equivalent to the estimated energy density of stellar radiation.

Eddington then attempts to specify a model for the spectrum of his estimated interstellar radiation field, based on his hypothesis of the statistical properties of stellar radiation:

> Radiation in interstellar space is about as far from thermodynamical equilibrium as it is possible to imagine, and although its density corresponds to 3.18° it is much richer in high-frequency constituents than equilibrium radiation of that temperature. [6, p. 371]

On this count, Eddington strayed from the data and that part of his analysis missed the mark.

The near-equality of Eddington's blackbody temperature of space and the CMB is considered a coincidence as "[t]he starlight radiation field is concentrated in galaxies like the Milky Way, which only occupy one part per million of the volume of the Universe, while the CMB fills the entire Universe" [7]. This comment demonstrates exactly the point raised in this Letter, and as we have been pointing out, it is hard for cosmologists to think outside of the CMB paradigm.

We also note several other non-cosmological estimates of the temperature of interstellar space that predate and that were more accurate than the Cosmic Microwave Background (CMB) Big Bang estimates [4]. Regener [8] predicted a value of 2.8 K in 1933 based on an analysis of the energy of cosmic rays arriving on Earth. This is remarkably close to the current best estimate of the value of a thermal blackbody spectrum at a temperature of 2.72548±0.00057 K [9]. Mackellar, following his identification of interstellar molecules [10], obtained the value 2.3 K in 1941, using the levels of excitation of the cyanogen molecule (CN) in intergalactic space [11].

3 Cosmic Microwave Background anisotropy

The CMB (or MBR) is highly isotropic, to roughly one part in 100 000. The spectral radiance contains small anisotropies which vary with the size of the region under examination. This anisotropy requires its own analysis separate from this Letter [12–14].

Suffice to say that advanced digital signal processing is performed on the data (e.g. [15]). A dipole anisotropy caused by the velocity of the Sun of about 370 km/s towards the constellation Leo, as determined from the MBR, is first subtracted from the Doppler shift of the background radiation. The root mean square (RMS) variations of the remainder are only 18 μK [7]. This anisotropy is a characteristic of the Microwave Background Radiation, whether it is of galactic or cosmological origin. Occam's razor favours a galactic origin.

4 Discussion and conclusion

In this Letter, we have pointed out that there were several non-cosmological estimates of the blackbody temperature of interstellar space that predated and that were more accurate than the Cosmic Microwave Background (CMB) Big Bang estimates. They are disregarded and considered coincidental as they are not based on the cosmological Big Bang model. We note the importance of this question, as the energy requirements of the two different explanations (galactic vs cosmological) are substantially different. We also point out that the actual correct explanation can't be determined from the measurements done in our local neighbourhood inside the Milky Way.

Received on September 25, 2021

References

1. Penzias A. A. and Wilson R. W. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal Letters*, 1965, v. 142, 419–421.
2. Dicke R. H., Peebles P. J. E., Roll P. J. and Wilkinson D. T. Cosmic Black-Body Radiation. *Astrophysical Journal Letters*, 1965, v. 142, 414–419.
3. Assis A. K. T., Neves M. C. D. History of the 2.7 K Temperature Prior to Penzias and Wilson. *Apeiron*, 1995, v. 2 (3), 79–84.
4. Assis A. K. T., Neves M. C. D. The Redshift Revisited. *Astrophysics and Space Science*, 1995, v. 227, 13–24.
5. Eddington, A. S. Diffuse Matter in Interstellar Space. *Roy. Soc. Proc.*, 1926, v. 111, 424–456.
6. Eddington, A. S. The Internal Constitution of the Stars. Cambridge University Press, Cambridge, 1926/1988 reprint, Chapter 13, p. 371.
7. Wright E. L. Eddington's Temperature of Space. www.astro.ucla.edu/~wright/Eddington-T0.html, last modified 23 October 2006.
8. Regener E. Der energiestrom der ultrastrahlung. *Zeit. Phys.*, 1933, v. 80, 666–669.
9. Fixsen, D. J. The Temperature of the Cosmic Microwave Background. *The Astrophysical Journal*, 2009, v. 707 (2), 916–920. arXiv: astro-ph/0911.1955.
10. Mackellar A. Evidence for the Molecular Origin of Some Hitherto Unidentified Interstellar Lines. *Publications of the Astronomical Society of the Pacific*, 1940, v. 52 (307), 187.
11. Mackellar A. *Pub. Dom. Astrophys. Observatory, Victoria, BC*, 1941, v. 7, 251.
12. Wright E. L. Theoretical Overview of Cosmic Microwave Background Anisotropy. In: Freedman W. L., ed. Measuring and Modeling the Universe. Cambridge University Press, Cambridge, 2004, p. 291. arXiv: astro-ph/0305591.
13. Challinor A. CMB anisotropy science: a review. arXiv: astro-ph.CO/1210.6008.
14. Hoyle F., Burbidge G. and Narlikar J. V. A Different Approach to Cosmology, From a Static Universe through the Big Bang towards Reality. Cambridge University Press, Cambridge, 2000.
15. Hivon E., Górski K. M., Barth Netterfield C., Crill, B. P., Prunet S., Hansen F. MASTER of the CMB Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex CMB Data Sets. arXiv: astro-ph/0105302.

LETTERS TO PROGRESS IN PHYSICS

In Memoriam of Simon Shnoll (1930–2021)

Dmitri Rabounski
Puschino, Moscow Region, Russia. E-mail: rabounski@yahoo.com

Simon Shnoll passed away on September 11, 2021 being 91 years old. He was born on the vernal equinox, March 21, 1930, and was one of the few greatest biophysicists of the 20th century, as well as a very good open hearth person. I have known him closely for many years. Here I want to pay tribute to his memory and his outstanding achievements in science.

Simon Shnoll was born on March 21, 1930, the day of vernal equinox, in Moscow, the USSR. His father Eli Shnoll was a religious philosopher belonging to the Russian Orthodox Church (as well as a polyglot who was fluent in more than twenty languages). In 1933 his father was imprisoned for his faith, together with many hierarchs of the Orthodox Church, to the prisoner camp on the Solovki Islands, but after three years of jail, in 1936, he was released as hopeless ill then died shortly before World War II. His mother Faina was a school teacher. His parents were religious; they had four sons. On the contrary, Simon Shnoll grew up an atheist; he told me many times: "My father was a religious philosopher and was fluent in twenty languages, while I know only one language and do not believe in God, but I believe in the Saints who lived in the past and those who live among us."

Simon Shnoll spent his childhood in Kaluga, a city on the Oka River, 160 km south-west of Moscow, where his family lived in exile. Due to job restrictions, his mother was largely unemployed: his family survived on odd jobs. Simon started to work commencing his 6 years as a herdsman in the summer season. His family experienced hunger; his youngest brother died in baby age of hunger in Autumn, 1941, because his mother lost milk due to hunger. Simon Shnoll always told me that only a person who has survived many years of hunger and watched how people have no power to bury the bodies of their relatives can understand the divine aroma of a freshly cooked loaf of bread.

Simon Shnoll first visited the Puschino area being an 11-year-old boy, in the summer of 1941 in search of casual work in the fields on the northern (low) bank of the Oka River opposite Puschino (located on the southern bank, at an altitude of 190 meters). There was a rest house for the officers' school cadets and a beach. Simon decided to freshen up in the river with the cadets who were sunbathing on the beach, but got into a whirlpool that sucked him weakened by hunger, and he began to drown, screaming for help. The cadets, these athletic young men, pulled him out of the river by the hair and, looking at his very skinny body, fed him as best they could. All these cadets were killed in action three months later, in October, 1941, on the fields near Moscow, when, armed only with rifles and grenades, they tried to stop the German tank columns moving towards Moscow (but they did it). The bridge near Puschino across the Oka River connecting Moscow and Tula is named after them.

Simon Shnoll. Puschino, about 2005.

The battleline very quickly approached Moscow, and at the end of August 1941, when 11-year-old Simon was grazing cattle in the fields on the northern (low) bank of the Oka River, he suddenly saw a chain of small fountains of dust rising from the ground and quickly approaching him, and then he heard pops shots from above: it was a German fighter pilot who decided to "hunt" a boy among a herd of cows... Simon first ran in a zigzag among the cows, then stumbled, fell and waited for death, but the German pilot used up all his ammunition and flew away.

When Simon returned back to Kaluga, there was chaos due to approaching the battleline. In the early morning of October 12, 1941, a military commissar knocked the door of Shnoll's room. The commissar went around all the civilian residents in their borough and told them: "Go away, today the Germans will come and kill you all." He was kind enough to help Shnoll's mom and all her four children get to the train station and then put them on the last train heading towards Moscow. His prediction was prophetic: within 2.5 months from October 12 to December 30, 1941, when Kaluga (50,000 inhabitants) was under occupation, the Germans killed more than 20,000 people in the ghetto and POW camp, including the captured soldiers and officers of the Kaluga garrison. A few years later, Shnoll's former neighbour, who moved to Moscow after the liberation, told him and his mother that, on their street, German soldiers often shot pedestrians just for fun, using them as targets for shooting training. Those were grey-dressed soldiers of Wehrmacht, and not SS. It was a bloody bacchanalia of war…

The refugees, including Shnoll's family were dropped off a train in a rural area far from Moscow, then Shnoll's mother and four her sons wandered around the villages of the Moscow region for a month in the cold autumn, begging for alms from the peasants. His younger brother, a baby, starved to death during this time. When Shnoll's mother with her three surviving children reached Moscow, they were allowed to stay there, because her husband, an "enemy of the people", had already died, so the restrictions on his family were lifted. Only 2.5 years later, in 1944, she found a permanent job as a teacher in an orphanage, where she and her sons could live and dine with orphans — "children of war".

The wartime in Moscow was accompanied by a surf of street crime. Even the immediate executions at the crime scene by police patrols could not stop the street robberies and murder. Once criminal teenagers who were older than Simon beat him to a pulp; they knocked out all of his front teeth. After this incident, Simon decided to take his fate into his own hands: he took a basic self-defence course from one of sports teachers. This saved his life many years later, in 1956, when the security officer who controlled Shnoll's work at that time, being completely drunk, tried to shoot him due to personal animosity. This happened in the forest on a river bank, during one of the trips of the laboratory employees to wild nature, near the campfire, where they all drank medical alcohol diluted half-to-half with water (they had much medical alcohol in the laboratory for biological purposes, while the laboratory rabbits were a good addition to the dinner table in the conditions of a total deficit of food products in the stores of the USSR). When the others were legless from the alcohol they drank, Simon Shnoll refused to continue drinking with the security officer, which sparked an outburst of his aggression. He told Shnoll: "I will kill you, and then all the Jews". Then he pulled his pistol out of the holster, sent a bullet into the pistol chamber and tried to shoot Simon. However, Simon turned him over with a judo technique, turned his hand with the pistol under his chin and… did not fire. Simon explained to me that this self-defence technique against an armed person necessarily ends with a shot to head: the fighter does not think, but automatically performs all these sequential movements, including the final headshot, which is achieved by long term training. But — said Simon — something stopped him at the last moment and he did not take the mortal sin of murder upon his soul. Simon stunned the security officer with a blow and then, having unloaded his pistol, threw all the bullets into the river. The security officer had a lot of respect for Simon after this incident.

Prior 1944, Simon Shnoll had never visited school. In Moscow, in 1942–1944 he worked as an electrician's assistant boy. In the meantime, he was educated at home by his mother, who was a very educated person. As a result, in 1944, being a 14-year-old boy, Simon Shnoll passed the 9th grade exam and graduated from school in 2 seasons in 1946.

In the summer of 1946, being 16 years old, Simon Shnoll entered the Department of Biology of Moscow University. It was the first peaceful summer after World War II, when hundreds of thousands of demobilized young soldiers tried to enter in universities. The number of applications per one student seat reached several hundreds. To enter, you had to pass all the entrance exams fine. At one of the entrance exams, 16-year-old Simon Shnoll met the love of his life, Maria Kondrashova, who was then 18 years old. She told me how this happened. She and Simon got up from their desks in the exam room at the same time and gave their exam papers to the examiner. The examiner began to check their papers at the same time, while they stood next to his desk. She watched at Simon. He looked like a small, skinny chick. She told me that she immediately felt a strong desire to warm and feed Simon like a child… After checking their exam papers, the examiner looked at them and said: "Both of your exam papers are good enough, but his paper is much better!" She immediately said a reply, looking towards Simon: "What a buster!" Very soon after that day, they realized that they could not live without each other and remained together for 74 years until she passed away on June 11, 2020. They had a son and daughter, as well as many grandchildren.

Maria Kondrashova was a biochemist. She and Simon were graduated from the same Department of Biology, where they had the same teacher, Prof. Sergey Severin, who introduced them into science. Maria and Simon lived in a small room in a shared apartment with neighbours in Moscow until 1963, when they were invited to live and work in Puschino, to a new research institute called the Institute of Theoretical and Experimental Biophysics (a.k.a. the Institute of Biophysics), where each of them headed his own research laboratory. Like Simon Shnoll, she first got a degree of Candidate of Sciences (which is analogous to the PhD degree), then — a degree of Doctor of Sciences. She was also a Professor of Moscow University. In addition to many of her other scientific works, her last major work, which she conducted since the 1990s, was a method of total diagnosis of the whole body using the analysis of just one drop of blood. Her outstanding scientific discoveries and ideas are scattered across many of her scientific articles in Russian, often in a very succinct concise form and, therefore, little known to the scientific community. These results and the technologies she developed are still awaiting rediscovery.

As for her personality, I would call her a "commander" in spite of the fact that she was a very nice and intelligent woman in everyday communication, in particular with me. She quickly recognized the identity of people and then behaved with them according to their personality. I discussed this issue with her. She explained to me that in science there is always only one leader who created an idea and developed it, and all the others are only assistants. Otherwise, without a "strong hand" any business, even a very good scientific development, would quickly collapse. Therefore, she always considered those who helped her in her laboratory to be mercenaries. In particular, therefore, she did not have followers in science. Science is moving forward by the forces of bright individuals, of whom there are very few, and not by scientific teams, as many mistakenly think, — she explained. Any scientific team shines with the reflected light emitted by its leader, and as soon as this leader ceases to exist, his or her team "goes out", ceasing to do something new, because only an individual, due to the strength of his of her will, is able to generate ideas and do scientific developments.

As for Simon, her husband, she told me that he is not really the "nice grey-haired old man" I have known for the past three decades. In fact, he is a very determined and risk-taking person, capable of, for example, surviving alone in a wild forest or repelling a gang of armed robbers. His psychology — she said — is rather similar to the psychology of a 14-year-old teenager, since only teenagers are not afraid of death and do everything without thinking about the consequences. The advantage of such a psychology is that Simon Shnoll had his own, completely independent point of view on many familiar things both in everyday life and in science.

Let us go back to Simon Shnoll and his story. After graduating from the Department of Biology in 1951, Simon Shnoll was under risk to be sent as a private for 5 years of military service (like his elder brother, a mathematician, who served 5 years in Red Army after the graduation). It was possible to be free from military service if he was hired by an institution where positions were equated to military service. After two months of nervous searches, Simon Shnoll took a job in a research laboratory of the Department of Radiology of the Institute for the Professional Development of Physicians. The laboratory was directly subordinated to the Soviet atomic project and included both military and civilian doctors; they studied how high doses of radiation exposure affect animals.

Despite the high level of secrecy in the laboratory, there was a complete mess and irresponsibility with safety measures for the staff during the radiation exposure of test animals. As a result, one day Simon Shnoll got a lethal dose of radiation. The female doctor who examined him said he would definitely die from radiation sickness within one month and all she could do for him is to give him as much pain reliever as needed. It was a very sad month in his life. He was ready to die, but his body had overcame the sickness in some incredible way. Moreover, he told me that his body had fully recovered without any repercussions, including the reproductive function. He explained that radiation exposure may not always be fatal, while radiation poisoning is always fatal due to the small radioactive particles penetrated into the body and permanently exposing it with their radiation.

Following this incident, Simon Shnoll became unfit for military service, but he continued to work in the Department of Radiology of the Institute for the Professional Development of Physicians. Then, since 1959, he began to lecture on biochemistry at the Physics Department of Moscow University, where he began as an associate professor, and then a professor until his death.

In 1963, Prof. Gleb Frank, a biophysicist, a member of the USSR Academy of Sciences, as well as a very influential organizer of science in the USSR, invited Simon and Maria to continue their scientific research in Puschino, a new "academic" town just erected 100 km south of Moscow on the southern (high) bank of the Oka River near the radio astronomical observatory of the Academy of Sciences. There were no "outsiders" in the new "academic" town: the peasants of the small village, Puschino, which had been located on the site for the past 600 or 800 years, were deported (except only a few persons who escaped the deportation) just before the construction of the town began in 1959. The absence of "outsiders" in the town created a unique social environment consisting only of scientists (and a small number of builders). Gleb Frank provided Simon and Maria with personal laboratories at a new research institute called the Institute of Biophysics, where he was Director. He also conversated with the town administration about providing them with a 4-bedroom apartment on the 9th (upper) floor of a just erected residential tower, which was luxurious living conditions compared to the small room in the shared apartment, where they huddled in Moscow. These were the first two "settlers" that Frank invited to live and work in Puschino. Simon always told me that Maria was the "settler number 1", and he was "number 2". On the contrary, Maria told me that he was the "settler number 1", and she was only "number 2".

For Simon Shnoll, it was a return to the areas of his childhood on the Oka riverbanks. Simon and Maria moved to Puschino with Maria's mother, who volunteered to took care of their son and daughter, while Simon and Maria spent all their time in their laboratories at the Institute. These were the times of the USSR, when workers received salary regardless of the results of their labour. Therefore, the quality of building work was low. The very first rain revealed many cracks in the waterproofing layer of the roof above the apartment, where Simon and Maria lived. Since there was nowhere to wait for help, Simon made a fire in front of the house, on which he melted down more than a dozen buckets of bitumen, and then going upstairs to the roof with the buckets of molten bitumen (the elevator in their house had not yet worked), filled all the cracks in the waterproofing layer. Also, the common heating system in the town sometimes malfunctioned during the winter seasons so that their slippers and tights of their little children froze to the floor. Nevertheless, life improved year after year. Fresh air (as opposed to the air in Moscow), a view of the endless Russian expanses and of a natural reserve on the opposite northern (low) bank of the Oka River (the bison natural reserve), as well as weekly picnics in the near forest neutralized all troubles of their first years in Puschino. A few years later, Simon and Maria hired workers who built a house for them in a cottage village of biophysicists near Puschino. Simon got carried away with the cultivation of an orchard and growing potatoes, but soon abandoned this "agricultural hobby", when realized that it takes away almost all of his time and is incompatible with his scientific research.

Simon travelled to Moscow twice a week, where he lectured at the Physics Department of Moscow University. Then there was no rapid express bus connecting Moscow and Puschino and private cab services (as now), and, therefore, such a trip took many hours. Also, in Puschino, as almost elsewhere in the USSR, there was a shortage of food products in stores; even with enough money, people could not freely buy what they wanted. In Puschino, this situation was solved by the fact that each Institute had its own restaurant, where employees had breakfasts and dinners with all their family members. However, the supply in the Moscow stores was good. Therefore, Simon usually travelled to Moscow on the eve of his lecture day. Arriving in Moscow in the evening, he usually purchased two full carry bags and a full backpack of food, then spent the night at the Facility of Biophysics of the Physics Department. Then, in the morning and in the daytime, he gave lectures, after which, already in the evening, he returned by bus to Puschino.

The life and all scientific achievements of Simon Shnoll and Maria Kondrashova after 1963 and until the last years of their lives were associated with Puschino and the Institute of Biophysics, which they considered their home. Shnoll's laboratory was one of the largest at the Institute: he told me that at the years of rise, he had 67 employees who did what he said. Since 1963, the staff of his laboratory were bestowed 26 Doctor of Sciences degrees and more than 120 Candidate of Sciences (PhD) degrees. In addition, many other laboratories of the Institute were founded by his employees, who decided to start their own research in another field of biophysics.

However, the years often change people or, most likely, as Jack London noticed, we often do not see the hidden character traits of some persons, which then become apparent over the years. In the beginning, Shnoll's laboratory staff and their families usually went out on a joint picnic at a weekend in the nearby forest. Then this picnic "decayed" into several smaller picnics, the participants of which tried to ignore others. Then, 10 years after the founding of the Institute, some of his former employees stopped greeting him, meeting him in the town or in the Institute… Maria told him that these were not true scientists, but those who simply wanted to "get well" in life; they got everything they wanted from him — scientific degrees and individual laboratories — and now they no longer need him. Simon Shnoll told me that this poor fact deeply hurt his heart, as he considered all his former employees to be his friends. During the mass "exodus" of scientists from Puschino in 1989–1991, just after the Iron Curtain that separated the USSR from the rest of the world has rushed down, many scientists left Puschino for the USA and Germany. According to Maria, after those two years some Institutes had become 75% empty. Their daughter had already lived in Moscow. Their son left for the USA with his wife and their children (they lived with Simon and Maria in the same 4-bedroom apartment): Simon and Maria were left alone in their apartment… Their circle of close friends narrowed even more; among them were remaining Eugeny Maevski and Heindrich Ivanitski, who held heading positions at their Institute and always supported Simon and Maria.

Simon Shnoll at the dinner table in his laboratory.
Puschino, 2005.

However, it was in the early 1990s that Simon Shnoll's research got the most rise (from the use of personal computers). To understand this, it is necessary to go back to 1951, when he first drew attention to a phenomenon that much later, in 2005, I called the *Shnoll effect*.

In 1951, Simon Shnoll, who had just graduated from university, began working in a research radiological laboratory subordinated to the Soviet atomic project (see above). Among other things, he conducted experiments to study the interaction of radioactive amino acids with muscle proteins (this was the topic of his PhD thesis). He discovered that the rate of this (very stable) chemical reaction, measured in hundreds of consecutive very accurate measurements taken during one working day through the same short time intervals, has systematic deviations from the average numerical value, which are not dependent of the experimental conditions, but only on the local time of measurement. It looked as if some very fine noise, with its repeated minima and maxima, was superimposed on the measurements of the very stable rate of this chemical reaction. The study of this systematic noise, its fine structure and origin became the main scientific task of Simon Shnoll throughout his life.

Continuing these studies in the 1950s in Moscow and then in the 1960–1970s in Puschino, Shnoll found that this specific noise is present not only in measurements of the rate of chemical reactions in muscle proteins, but also in any proteins in general. Moreover, in the 1970s, he found this noise in any biochemical reactions that he tested. Even more, in 1970–1980s, he found this noise, with its specific minima and maxima, in any consecutive physical measurement that he tested or analysed, unnecessary biochemical reactions. For example, he found it by measuring the rate of alpha-decay and beta-decay of atomic nuclei. In general, he found the following: the more stable the signal was, the better this noise manifested itself.

When I first met Simon Shnoll in 1991, I asked him: "What should be measured to detect this noise and its fine structure?" He adjusted his glasses with his hand and answered me: "It does not matter what!"

It should be noted that registering the systematic noise in an experiment is only a small percentage of the whole problem. The main trouble is the processing of the measurement results, which allows you to "see" this noise and its fine structure. In the pre-computer era, when performing these experiments based on a sequence of measurements of a signal for one day, month, or year, you had to do the following. First, you had to manually create on paper a histogram of the measured numerical values of the signal for each measurement interval, say, for each interval of 15 or 30 seconds. This would result in about 2,880 hand-drawn histograms per day, about 86,400 hand-drawn histograms per month, and over 1 million hand-drawn histograms per year for 30-second measurement intervals (double the number of histograms for 15-second intervals). Then you had to visually compare all these histograms with each other to find the ones that match each other in their shape. And finally, you had to create a graph that shows when the found similar histograms appear according on the local time. In the pre-computer era, Shnoll was forced to limit himself to only analysing the results of measurements obtained within one or two days. Even so, he had to create over 3,000 hand-drawn histograms for each experiment and then visually compare all those hand-drawn histograms to each other. This is clearly an overwhelming task for one person. Dozens of his employees, mostly young women who graduated from Moscow University, drew these thousands of histograms by hand every day and then compared them with each other. It was a titanic work!

Things got much easier in 1997, when Edwin Pozharski, a young man from Poland, who, just graduated from Physical Technical Institute that is to north of Moscow (he was engaged in X-ray analysis of proteins), created a computer program allowing to create and analyse dozens and hundreds of thousands of histograms. In particular, the use of his program allowed to create and analyse histograms for measurements performed over a week, month and even a year. He created this program not for fame or money (he did it on a volunteer basis), but simply out of great respect for Shnoll and his research. His program has gone through many updates and is still the main working tool in the study of the Shnoll effect. Thanks to him!

The next 10 years of Shnoll's research after 1997 were the most fruitful. Using the computer program to create histograms and analyse them, he found that the discovered fine structure is manifested in any random noise, and not only in the random noises registered in biological and physical processes. In particular, he found the same fine structure in the random noise generated by a random number generator on a computer.

To summarize briefly the Schnoll effect, it is best to give a quote from my 2014 article*, where I already did it:

> "The principal error in understanding the Shnoll effect is that some people think it is a periodical fluctuation of the magnitude of the signal that is measured. This is incorrect, since the magnitude of the signal and the average noise remain the same during the long-term measurements done by Shnoll and his workgroup. Further, such processes are specifically chosen for the study that are very stable in time. Simply put, nothing allegedly changes in the experiments which continue during days, months, and even years. The subject of the measurement is the *fine structure of the noise* registered in stable processes.
>
> Every process contains noise. The noise originates due to the influence of random factors and satisfies the Gaussian distribution (i.e., the Gauss continuous distribution function of the probability of the measured value between any two moments of time). Gaussian distribution is attributed to any random process, such as noise, and is based on the averaging and smoothing of the noise fluctuation measured during a long enough interval of time. Nevertheless, if considering very small intervals of time, the real noise has a bizarre structure of the probability distribution function, which differs for each interval of time. Each of these real functions being considered "per se" cannot be averaged to a Gaussian curve. This is what Shnoll called the fine structure of noise and is the object of research studies originally conducted by Simon Shnoll, commencing in 1951–1954 to this day.
>
> So, the magnitude of noise is measured in a very stable process during a long enough duration of time (days, months, and even years). Then the full row of the measured data is taken under study. The full duration of time is split into small intervals. A histogram of the probability distribution function is then created for each of the small intervals. Each interval of time has

---

*Rabounski D. and Borissova L. General relativity theory explains the Shnoll effect and makes possible forecasting earthquakes and weather cataclysms. *Progress in Physics*, 2014, v. 10, issue 2, p. 63–70.

its own bizarre distribution function (form of the histogram) that differs from Gaussian function. Nevertheless, Shnoll found that "paired histograms," which have a very similar (almost identical) form, exist along the row of the measured data. That is, the histogram created for each interval of time has its own "twin" which has a similar form. The similar form was found in the histograms which were registered with the following periods of repetition connected with stars, the Sun, and the Moon:

- 24 hours = 1440 min (solar day);
- 23 hours, 56 min = 1436 min (stellar day);
- 24 hours, 50 min = 1490 min (lunar day);
- 27 days, 7 hours, 43 min = 39 343 min (lunar month);
- 31 days, 19 hours, 29 min = 45 809 min (period of the lunar evection);
- 365 days = 525 600 min (calendar year);
- 365 days, 6 hours, 9 min = 525 969 min (stellar year).

Also, aside as the similar forms of histograms, appearance the mirrored forms of histograms was registered by Shnoll with periods of:

- 720 min (half of the calendar/solar day);
- 182 days, 12 hours = 262 800 min (half of the calendar/solar year).

Shnoll called this phenomenon the "palindrome effect". It is one of Shnoll's newest findings: despite his having started the research studies in 1951, the possibility of the appearance of the mirrored forms of histograms only came to his attention in 2004. The "palindrome effect" was first registered in December 2007. Aside from these two periods of the "palindromes", a number of other palindrome cycles were found. However, certain circumstances have not allowed a continuation of these studies in full force yet.

As was shown by Shnoll after many experiments done synchronously at different locations from South Pole to North Pole, an appearance of the similar form (or the mirrored form) of the histograms does not depend on the geographical latitude, but depends only on the geographical longitude, i.e., the same *local time* at the point of observation. In other words, the Shnoll effect is manifested equally at any location on the Earth's surface, according to the local time, meaning the same locations of the celestial objects in the sky with respect to the visible horizon.

It is significant that the process producing the noise that we measure can be absolutely anything. Initially, in 1951, Shnoll started his research studies from measurements of the speed of chemical reactions in the aqueous solutions of proteins. Then many other biochemical processes attracted his attention. After decades of successful findings, he focused on such purely physical processes as alpha-decay and beta-decay of the atomic nuclei. It was shown that not only all the random natural processes of different origins, but even artificial processes as random-number generation by computer software manifest the Shnoll effect. In other words, this is a fundamental effect."

Shnoll told me that neither signal level nor noise level is actually measured in his experiments: their numerical values remain very stable during measurement. Only standard time intervals between adjacent measurements change with the periods that he registered. That is, signal level and noise level remain unchanged, while the "unit time interval" between adjacent measurements is not "unit" but changes according to the fine inhomogeneous structure of space-time, through which the observer, together with his laboratory and the Earth itself, travels in the cosmos. The observer's laboratory is located somewhere on the surface of the Earth, while the Earth's body revolves around its axis, the Earth revolves around the Sun, the planets revolve around the Sun, and the entire Solar System travels along its specific trajectory in the Galaxy… As a result, the observer together with his laboratory travels in the cosmos through the fine structured grid of the standard "unit time intervals", which is caused by the fields of the aforementioned rotating cosmic bodies. This motion of the observer leads to the fact that his measurements of everything are affected by the corresponding periodic changes in the duration of the standard "unit time interval" between consecutive measurements. The more stable the quantity that he measures, the more obvious the fine structure of the grid of time intervals through which he moves in the cosmos.

In other words, Shnoll believed that the fine structure of random noise discovered by him (a.k.a. the Shnoll effect) manifests the fine inhomogeneous structure of time itself[*].

In 2007–2008, I was honoured to be the editor of Shnoll's book, in which he explained the entire story of his discovery, starting in 1951, as well as all the details of his experiments and experimental results obtained up to those years. Prior that time, his experimental results were scattered over many dozens of his fragmentary papers. He asked me to help him with the structure of the book: he drafted many chapters on different topics that needed to be somehow connected with each other and combined into a whole book. We spent many hours together discussing every detail of the book. The book was published in Russian in 2009, and its English translation in 2012.[†] To be honest, I was should translate his book myself, because I knew the subject of the book like no one else. But then I was so physically exhausted that Maria Kondrashova took pity on me and invited two women for translation. Now, I have a great regret about this missed opportunity.

---

[*] Shnoll S.E. Changes in the fine structure of stochastic distributions as a consequence of space-time fluctuations. *Progress in Physics*, 2006, v. 2, issue 2, p. 39–45.

[†] Shnoll S. E. Cosmophysical Factors in Stochastic Processes. American Research Press, Rehoboth (NM), 2012, 388 pages.

This obituary turned out to be very personal, far from officialdom...

I, like most residents of Puschino, often walk along the Green Zone, the main street of the town, which is a 1,700-meter forested boulevard that runs through the town and separates the Institutes from the residential area. This boulevard is framed by the Avenue of Science from the side of the residential buildings and by the Institute Avenue from the side of the Institutes. At the entrance of each of the Institutes, I see memorial plaques on the wall dedicated to the influential scientists of the Soviet period, members of the USSR Academy of Sciences, who worked in Puschino. In addition, several streets of the town are named after some of them. These influential people of the Soviet period were successful organizers of science, rather than outstanding researchers. Their scientific achievements were particular and had not changed biophysics or biochemistry as a whole, while the technologies they have developed (like any technologies in general) are rather the subject of industrial corporations than science: the task of scientists is to discover fundamental laws, which industrial corporations, if they deem necessary, can then use then to develop some technologies.

On the contrary, the Shnoll effect is a fundamental discovery. Understanding this fine structure of the pattern of time through which we, together with the planet Earth, travel in the cosmos, will undoubtedly change not only biophysics as a science and physics in general, but our entire understanding of the world. In this sense, Simon Shnoll is similar to Copernicus, who also once changed the entire understanding of the world. I therefore have no doubt that once the scientific community has evolved enough to understand the significance of the Shnoll effect, then the Green Zone, the main boulevard that runs through the entire town, will be renamed Shnoll Boulevard. This will be not only fair, because Simon Shnol will forever remain the most outstanding research scientist who lived and worked in Puschino, but also symbolic — Shnoll Boulevard, running as a narrow through the entire town of scientists.

Let his memory live for ever!

Received on September 28, 2021

Physics of Transcendental Numbers on the Origin of Astrogeophysical Cycles

Hartmut Müller

Rome, Italy.
E-mail: hm@interscalar.com

Transcendental ratios of physical quantities can provide stability in complex dynamic systems because they inhibit the occurrence of destabilizing resonance between the elements of the system. In this paper we analyze recently discovered astrophysical and geophysical cycles in order to verify this numeric-physical paradigm.

Introduction

The Solar system behaves like a precise chronometer. Indeed, the orbital and rotational periods of the planets, planetoids and large moons are exceptionally stable. In view of the huge number (more than 800.000) of orbiting and rotating celestial bodies, perturbation models based on conventional theories of gravitation predict long-term highly unstable states [1, 2] and have a problem with the real stability of the Solar system. Moreover, they do not explain basic facts, for instance, why the Solar system has established the orbital periods 90560 days (Pluto), 60182 (Neptune), 30689 (Uranus), 10759 (Saturn), 4333 (Jupiter), 1682 (Ceres), 687 (Mars), 365 (Earth), 225 (Venus) and 88 days (Mercury). The current distribution of the planetary and lunar orbital and rotational periods appears to them to be completely coincidental.

Recently discovered astrophysical and geophysical cycles of galactic origin suggest that despite the huge number of stars (more than 200 billion), our Galaxy behaves like a precise chronometer as well. Disappointingly, there is no theory of gravitation that derives the correct movement of stars in galaxies or explains at least the existence of galaxies without introducing a huge amount (corrently 68%) of dark energy [3]. In spiral galaxies, the orbiting of stars seems to strongly disobey both Newton's law of universal gravitation and general relativity. Recently, an 85% dark matter [4] universe is required for saving the conventional paradigm.

Perhaps the concept of gravitation itself requires a revision. Obviously, it is not about details, but an important part of the hole is missing.

In this paper we introduce a basic numeric-physical approach that could be the missing link as it allows resolving stability tasks in dynamic systems of any level of complexity.

Methods

In [5] we have shown that the difference between rational, irrational algebraic and transcendental numbers is not only a mathematical task, but it is also an essential aspect of stability in complex dynamic systems. For instance, integer frequency ratios provide resonance interaction that can destabilize a system [6]. Actually, it is transcendental numbers that define the preferred ratios of quantities which avoid destabilizing resonance interaction [7]. In this way, transcendental ratios of quantities sustain the lasting stability of periodic processes in complex dynamic systems. With reference to the evolution of a planetary system and its stability, we may therefore expect that the ratio of any two orbital periods should finally approximate a transcendental number [8].

Among all transcendental numbers, Euler's number $e = 2.71828\ldots$ is unique, because its real power function e^x coincides with its own derivatives. In the consequence, Euler's number allows inhibiting resonance interaction regarding any interacting periodic processes and their derivatives. Because of this unique property of Euler's number, complex dynamic systems tend to establish relations of quantities that coincide with values of the natural exponential function e^x for integer and rational exponents x.

Therefore, we expect that periodic processes in real systems prefer frequency ratios close to Euler's number and its rational powers. Consequently, the logarithms of their frequency ratios should be close to integer $0, \pm 1, \pm 2, \ldots$ or rational values $\pm 1/2, \pm 1/3, \pm 1/4, \ldots$ In [9] we exemplified our hypothesis in particle physics, astrophysics, cosmology, geophysics, biophysics and engineering.

Based on this hypothesis, we introduced a fractal model of matter [10] as a chain system of harmonic quantum oscillators and could show the evidence of this model for all known hadrons, mesons, leptons and bosons as well. In [11] we have shown that the set of stable eigenstates in such systems is fractal and can be described by finite continued fractions:

$$\mathcal{F}_{jk} = \ln(\omega_{jk}/\omega_{00}) = \langle n_{j0}; n_{j1}, n_{j2}, \ldots, n_{jk} \rangle \qquad (1)$$

where ω_{jk} is the set of angular eigenfrequencies and ω_{00} is the fundamental frequency of the set. The denominators are integer: $n_{j0}, n_{j1}, n_{j2}, \ldots, n_{jk} \in \mathrm{Z}$. The cardinality $j \in \mathrm{N}$ of the set and the number $k \in \mathrm{N}$ of layers are finite. The numeric occupancy of one layer does not influence the numeric occupancy of other layers, so that each layer can be considered as an independent dimension of a k-dimensional manifold. In the canonical form, all numerators equal 1. We use angle brackets for continued fractions.

Any finite continued fraction represents a rational number [12]. Therefore, the ratios ω_{jk}/ω_{00} of eigenfrequencies are always irrational, because for rational exponents the natural exponential function is transcendental [13]. This circumstance provides for lasting stability of those eigenstates of a chain system of harmonic oscillators because it prevents resonance interaction [14] between the elements of the system. The distribution density of stable eigenstates reaches local maxima near reciprocal integers $\pm 1/2, \pm 1/3, \pm 1/4, \ldots$ that are attractor points (fig. 1) in the fractal set $\mathcal{F}_{jk}$ of natural logarithms. Integer logarithms $0, \pm 1, \pm 2, \ldots$ represent the most stable eigenstates (main attractors).

In the case of harmonic quantum oscillators, the continued fractions $\mathcal{F}_{jk}$ define not only fractal sets of natural angular frequencies ω_{jk}, angular accelerations $a_{jk} = c \cdot \omega_{jk}$, oscillation periods $\tau_{jk} = 1/\omega_{jk}$ and wavelengths $\lambda_{jk} = c/\omega_{jk}$ of the chain system, but also fractal sets of energies $\mathrm{E}_{jk} = \hbar \cdot \omega_{jk}$ and masses $\mathrm{m}_{jk} = \mathrm{E}_{jk}/c^2$ which correspond with the eigenstates of the system. For this reason, we call the continued fraction $\mathcal{F}_{jk}$ the *Fundamental Fractal* of stable eigenstates in chain systems of harmonic quantum oscillators.

Fig. 1: The distribution of stable eigenvalues of $\mathcal{F}_{jk}$ for $k = 1$ (above) and for $k = 2$ (below) in the range $-1 \leqslant \mathcal{F}_{jk} \leqslant 1$.

In fact, scale relations in particle- [10] and astrophysics [15] obey the same Fundamental Fractal (1), without any additional or particular settings. The proton-to-electron rest energy ratio approximates the first layer of the Fundamental Fractal that could explain their exceptional stability. In fact, the life-spans of the proton and electron top everything that is measurable, exceeding 10^{29} years [16].

Property	Electron	Proton
$\mathrm{E} = \mathrm{mc}^2$	0.5109989461(31) MeV	938.2720813(58) MeV
$\omega = \mathrm{E}/\hbar$	$7.76344 \cdot 10^{20}$ Hz	$1.42549 \cdot 10^{24}$ Hz
$\tau = 1/\omega$	$1.28809 \cdot 10^{-21}$ s	$7.01515 \cdot 10^{-25}$ s
$\lambda = c/\omega$	$3.86159 \cdot 10^{-13}$ m	$2.10309 \cdot 10^{-16}$ m

Table 1: The basic set of the physical properties of the electron and proton. Data from Particle Data Group [16]. Frequencies, oscillation periods and wavelengths are calculated.

The proton-to-electron ratio (tab. 1) approximates the seventh power of Euler's number and its square root:

$$\ln\left(\frac{\omega_p}{\omega_e}\right) = \ln\left(\frac{1.42549 \cdot 10^{24}\,\mathrm{Hz}}{7.76344 \cdot 10^{20}\,\mathrm{Hz}}\right) \simeq 7 + \frac{1}{2} = \langle 7; 2 \rangle$$

In the consequence of this potential difference of the proton relative to the electron, the scaling factor $\sqrt{e} = 1.64872\ldots$ connects attractors of proton stability with similar attractors of electron stability in alternating sequence.

These unique properties of the electron and proton predestinate their physical characteristics as fundamental units. Table 1 shows the basic set of electron and proton units that can be considered as a fundamental metrology (c is the speed of light in a vacuum, $\hbar$ is the Planck constant). In [11] was shown that the fundamental metrology (tab. 1) is completely compatible with Planck units [17]. Originally proposed in 1899 by Max Planck, these units are also known as natural units, because the origin of their definition comes only from properties of nature and not from any human construct. Max Planck wrote [18] that these units, "regardless of any particular bodies or substances, retain their importance for all times and for all cultures, including alien and non-human, and can therefore be called natural units of measurement". Planck units reflect the characteristics of space-time.

We assume that scale invariance according to the Fundamental Fractal (1), which is calibrated to the physical properties of the proton and the electron, is a universal characteristic of organized matter and criterion of stability. This hypothesis we have called *Global Scaling* [9].

In [19] we applied the Fundamental Fractal (1) to macroscopic scales interpreting gravity as quantum attractor effect of its stable eigenstates. We have shown that the orbital and rotational periods of planets, planetoids and large moons of the solar system correspond with attractors of electron and proton stability [11]. This is valid also for exoplanets [15] of the systems Trappist 1 and Kepler 20. In [8] we have shown that the maxima in the frequency distribution of the orbital periods of 1430 exoplanets listed in [20] correspond with attractors of the Fundamental Fractal. In [21] we have shown that the maxima in the frequency distribution of the number of stars in the solar neighborhood as function of the distance between them correspond well with attractors of the Fundamental Fractal.

In this paper we will show that the Fundamental Fractal (1) determines also the Earth axial precession cycle, the obliquity variation cycle as well as the apsidal precession cycle and the orbital eccentricity cycle. In addition, we will show that recently discovered geological cycles, as well as the periodic variations in the movement of the Solar system through the Galaxy, substantiate their determination by the Fundamental Fractal.

Results

Since its birth the Sun has made about 20 cycles around the Galaxy, and during this time the Solar system has made many passages through the spiral arms of the disk. The Sun's orbit in the Galaxy is not circular. There are temporal variations in the distance from the Galactic center with a period of T_S = 170 million years [22] that corresponds precisely with the main attractor $\langle 90 \rangle$ of proton stability of the Fundamental Fractal (1):

$$\ln\left(\frac{T_S}{2\pi \cdot \tau_p}\right) = 90$$

$2\pi \cdot \tau_p$ is the oscillation period of the proton (tab. 1). The recently [23] discovered geological cycle with a period of $T_G = 27$ million years corresponds well with the same attractor ⟨90⟩, but relative to the angular oscillation period of the proton:

$$\ln\left(\frac{T_G}{\tau_p}\right) = 90$$

The connection $T_S = 2\pi \cdot T_G$ suggests that the 27 million years' geological cycle could be caused by angular components of the periodical variations of the distance of the Solar system (and the Earth) from the Galactic center. In addition, [23] reports a geological cycle of 8.9 Ma that approximates the main attractor ⟨87⟩ of proton stability:

$$\ln\left(\frac{8.9\ Ma}{2\pi \cdot \tau_p}\right) = 87$$

The Sun's path oscillates above and below the Galactic plane with a period of approximately 63 million years [22] that coincides with the main attractor ⟨89⟩ of proton stability:

$$\ln\left(\frac{63\ Ma}{2\pi \cdot \tau_p}\right) = 89$$

Earth's axial precession cycle (25,770 years) fits the attractor ⟨83⟩ of proton stability:

$$\ln\left(\frac{25,770\ a}{\tau_p}\right) = 83$$

By the way, 25,770 years is also the time it takes for a signal to travel from the Galactic center to Earth at the speed of light.

The Fundamental Fractal (1) is of pure numeric origin, and there is no particular physical mechanism that creates it. It is all about transcendental ratios of frequencies [8] that inhibit destabilizing resonance interaction. In this way, the Fundamental Fractal concerns all repetitive processes, independently on their temporal or spatial scales.

For instance, Earth's apsidal precession cycle and orbital eccentricity cycle (both of 112,000 years) correspond with the attractor ⟨77⟩ of electron stability:

$$\ln\left(\frac{112,000\ a}{\tau_e}\right) = 77$$

τ_e is the angular oscillation period of the electron (tab. 1). Earth's obliquity variation cycle (41,000 years) corresponds with the attractor ⟨76⟩ of electron stability:

$$\ln\left(\frac{41,000\ a}{\tau_e}\right) = 76$$

Naturally, we expect the existence of further galactic cycles that correspond with other main attractors of the Fundamental Fractal. Table 2 gives an overview of expected main attractor cycles in the scale of millions of years.

n	$T_p(n)$, Ma	$t_p(n)$, Ma	n	$T_e(n)$, Ma	$t_e(n)$, Ma
91	463.35	73.75	83	285.41	45.42
90	170.46	27.13	82	105.00	16.71
89	62.71	9.98	81	38.62	6.15
88	23.07	3.67	80	14.21	2.26
87	8.49	1.35	79	5.23	0.83
86	3.12	0.50	78	1.92	0.31

Table 2: Cycles corresponding with main attractors of proton and electron stability in the range of millions of years (Ma).

Every attractor of proton or electron stability defines the period of a stable cycle and its angular period. As main attractors correspond with integer exponents n of the Fundamental Fractal (1), it is easy to calculate main attractor cycles:

$$t_e(n) = \tau_e \cdot e^n \qquad T_e(n) = 2\pi \cdot t_e(n)$$

$$t_p(n) = \tau_p \cdot e^n \qquad T_p(n) = 2\pi \cdot t_p(n)$$

In general, the identification of the predicted galactic cycles requires a significant increase in current data precision.

Conclusion

Within our approach, numeric attractors of stability determine the distribution of matter in space and time. Since the distribution of the attractors is fractal, the distribution of matter is also fractal. Numerical attractors cause effects known as gravity, electricity, magnetism, and nuclear forces. Numerical relationships are primary, physical effects are secondary. Numerical attractors cause the formation of matter in all scales – from the electron and proton up to planets, stars and galaxies. Interscalar cosmology [9] bases on this approach.

In particular, for maintaining stability of motion, the Sun does not have to avoid parametric resonance with every single other star on its path through the Galaxy. As this task cannot be resolved in general, the application of transcendental frequency ratios appears to be a significant alternative. As we have shown, not only stars [21], but also planets [8] make extensive use of it.

Acknowledgements

The author is grateful to Leili Khosravi, Ulrike Granögger, Michael Kauderer, Valery Kolombet, Oleg Kalinin, Viktor Bart and Viktor Panchelyuga for valuable discussions.

Eternal thanks go to the sadly deceased great scientist Simon Shnoll. He found empirical evidence of cosmophysical factors that underlie all processes in the universe. Simon Shnoll's work stimulated the development of the author's approach over many years.

Submitted on October 1, 2021

References

1. Heggie D. C. The Classical Gravitational N-Body Problem. arXiv: astro-ph/0503600v2, 11 Aug 2005.
2. Hayes B. The 100-Billion-Body Problem. *American Scientist*, v. 103, no. 2, 2015.
3. Li M. et al. Dark Energy. arXiv:1103.5870v6 [astro-ph.CO] 7 Oct 2011.
4. Einasto J. Dark Matter. arXiv:0901.0632v2 [astro-ph.CO] 19 Oct 2010.
5. Müller H. On the Cosmological Significance of Euler's Number. *Progress in Physics*, 2019, v. 15, 17–21.
6. Dombrowski K. Rational Numbers Distribution and Resonance. *Progress in Physics*, 2005, v. 1, no. 1, 65–67.
7. Müller H. The Physics of Transcendental Numbers. *Progress in Physics*, 2019, vol. 15, 148–155.
8. Müller H. Physics of Transcendental Numbers meets Gravitation. *Progress in Physics*, 2021, vol. 17, 83–92.
9. Müller H. Global Scaling. The Fundamentals of Interscalar Cosmology. *New Heritage Publishers*, Brooklyn, New York, USA, ISBN 978-0-9981894-0-6, (2018).
10. Müller H. Fractal Scaling Models of Natural Oscillations in Chain Systems and the Mass Distribution of Particles. *Progress in Physics*, 2010, v. 6, 61–66.
11. Müller H. Scale-Invariant Models of Natural Oscillations in Chain Systems and their Cosmological Significance. *Progress in Physics*, 2017, v. 13, 187–197.
12. Khintchine A.Ya. Continued fractions. University of Chicago Press, Chicago, (1964).
13. Hilbert D. Über die Transcendenz der Zahlen e und π. *Mathematische Annalen*, 1893, v. 43, 216–219.
14. Panchelyuga V.A., Panchelyuga M. S. Resonance and Fractals on the Real Numbers Set. *Progress in Physics*, 2012, v. 8, no. 4, 48–53.
15. Müller H. Global Scaling of Planetary Systems. *Progress in Physics*, 2018, v. 14, 99–105.
16. Tanabashi M. et al. (Particle Data Group), *Phys. Rev. D*, 98, 030001 (2018), www.pdg.lbl.gov
17. Astrophysical constants. Particle Data Group, pdg.lbl.gov
18. Planck M. Über Irreversible Strahlungsvorgänge. *Sitzungsbericht der Königlich Preußischen Akademie der Wissenschaften*, 1899, v. 1, 479–480.
19. Müller H. Gravity as Attractor Effect of Stability Nodes in Chain Systems of Harmonic Quantum Oscillators. *Progress in Physics*, 2018, v. 14, 19–23.
20. Catalog of Exoplanets. Observatoire de Paris, http://exoplanet.eu/
21. Müller H. Physics of Transcendental Numbers Determines Star Distribution. *Progress in Physics*, 2021, v. 17, 164–167.
22. Gies D. R., Helsel J. W. Ice age epochs and the Sun's path through the Galaxy. *The Astrophysical Journal*, 2005, v. 626, 844–848.
23. Rampino M. R., Caldeira K., Zhu Y. A pulse of the Earth: A 27.5-Myr underlying cycle in coordinated geological events over the last 260 Myr. *Geoscience Frontiers*, 2021, v. 12, 101245.

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