Issue 2 2019 | Volume 15

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

PROGRESS IN PHYSICS

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

ISSN 1555-5534

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

PROGRESS IN PHYSICS

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

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

Postal Address

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

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

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

July 2019 **Vol. 15, Issue 2**

CONTENTS

Information for Authors

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

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

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

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

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

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

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

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

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

The Dirac Equation and Its Relationship to the Fine Structure Constant According to the Planck Vacuum Theory

William C. Daywitt

National Institute for Standards and Technology (retired), Boulder, Colorado. E-mail: wcdaywitt@me.com

The Dirac equation and the fine structure constant are complementary and cannot be understood separately. The manifestly covariant Dirac equation in the Planck vacuum (PV) theory (8) is a coupling-charge equation, where e_*^2 is the squared coupling charge that couples the equation to the PV state. The laboratory-measured electron or proton mass is denoted by m. The corresponding fine structure constant is $\alpha \equiv e^2/e_*^2$ where e^2 is the squared charge of the electron or proton as measured in the laboratory. Both the Dirac particle spin and the fine structure constant have their origin in the electron or proton coupling to the PV state. The electron g-factor, with radiative corrections, is calculated from the fine structure constant; and the proton g-factor is roughly estimated from the electron g-factor and the proton structure constant. The radiative corrections in the QED theory are the result of photon interactions taking place within the *pervaded* PV state. The apparent ability of the electron to emit and absorb photons is due to the ability of the PV state to emit and absorb photons to and from free space.

1 Introduction

The theoretical foundation [1] [2] [3] of the PV theory rests upon the unification of the Einstein, Newton, and Coulomb superforces:

$$\frac{c^4}{G}\left(=\frac{m_*c^2}{r_*}\right)=\frac{m_*^2G}{r_*^2}=\frac{e_*^2}{r_*^2} \tag{1}$$

where the ratio c^4/G is the curvature superforce that appears in the Einstein field equations. G is Newton's gravitational constant, c is the speed of light, m_* and r_* are the Planck mass and length respectively [4, p.1234], and e_* is the coupling charge.

The two particle/PV coupling forces

$$F_e(r)=\frac{e_*^2}{r^2}-\frac{m_ec^2}{r} \quad \text{and} \quad F_p(r)=\frac{e_*^2}{r^2}-\frac{m_pc^2}{r} \tag{2}$$

the electron core $(-e_*, m_e)$ and proton core $(+e_*, m_p)$ exert on the invisible PV state; along with their coupling constants

$$F_e(r_e)=0 \quad \text{and} \quad F_p(r_p)=0 \tag{3}$$

and the resulting Compton radii

$$r_e=\frac{e_*^2}{m_ec^2} \quad \text{and} \quad r_p=\frac{e_*^2}{m_pc^2} \tag{4}$$

lead to the important string of Compton relations

$$r_em_ec^2=r_pm_pc^2=e_*^2=r_*m_*c^2 \quad (=c\hbar) \tag{5}$$

for the electron and proton cores, where $\hbar$ is the reduced Planck constant. The Planck particle Compton radius is $r_*=e_*^2/m_*c^2$, which is derived by equating the Einstein and Coulomb superforces from (1). To reiterate, the equations in (2) represent the forces the free electron or proton cores exert on the invisible PV space, a space that is itself pervaded by a degenerate collection of Planck-particle cores $(\pm e_*, m_*)$ [5]. The positron and antiproton cores are $(+e_*, m_e)$ and $(-e_*, m_p)$ respectively.

Finally, the Lorentz invariance of the coupling constants in (3) lead to the energy

$$i\hbar\frac{\partial}{\partial t}=ie_*^2\frac{\partial}{\partial ct} \tag{6}$$

and momentum

$$-i\hbar\nabla=-i\frac{e_*^2}{c}\nabla \tag{7}$$

operators of the quantum theory [5]. It should be noted that the two operators are proportional to the squared coupling charge e_*^2.

Section 2 expresses the Dirac equation in terms of PV parameters. Section 3 discusses the fine structure constant. Section 4 discusses the gyromagnetic g-factor. Section 5 discusses the electron g-factor and Section 6, the proton g-factor. Sections 5 and 6 are a work in progress that seek to relate the QED radiative corrections to the PV coupling model. Section 7 presents some comments and conclusions.

2 Dirac equation

Using (5), the manifestly covariant form [6, p.90] [Appendix A] of the Dirac equation for the Dirac particle cores (electron, positron, proton, antiproton) can be expressed as:

$$\left(ic\hbar\gamma^\mu\frac{\partial}{\partial x^\mu}-mc^2\right)\psi=\left(ie_*^2\gamma^\mu\frac{\partial}{\partial x^\mu}-mc^2\right)\psi= \tag{8}$$

$$\left[ie_*^2\gamma^0\frac{\partial}{\partial x^0}+i\begin{pmatrix}0 & cS_j\\ -cS_j & 0\end{pmatrix}\frac{\partial}{\partial x^j}-mc^2\right]\psi=0 \tag{9}$$

where the second term in (9) is summed over $j = 1, 2, 3$ and

$$\begin{pmatrix} 0 & cS_j \\ -cS_j & 0 \end{pmatrix} = \begin{pmatrix} 0 & e_*^2\sigma_j \\ -e_*^2\sigma_j & 0 \end{pmatrix} \tag{10}$$

where one of the charges in e_*^2 belongs to the free particle and the other to any one of the Planck-particle cores within the degenerate PV state. The $e_*^2\sigma_j/c$ from the 4x4 matrix in (10) are the 2x2 spin components of the S-vector

$$\vec{S} = \frac{e_*^2}{c}\vec{\sigma} \quad \left(= \hbar\vec{\sigma}\right) \tag{11}$$

that applies to all the Dirac particles. $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is the Pauli spin vector, where the σ_js are 2x2 matrices.

3 Fine structure constant

Using the expressions in (5), the fine structure constant can be expressed as

$$\alpha = \frac{e^2}{e_*^2} = \frac{e^2}{r_* m_* c^2} = \frac{e^2}{r_p m_p c^2} = \frac{e^2}{r_e m_e c^2} \tag{12}$$

where e is the magnitude of the laboratory-observed electron/proton charge. If $e = e_*$, then the Compton relations in (5) yield $\alpha = 1$ for the Dirac equation. Thus it is clear that the fine structure constant provides the "bridge" over which the Dirac equation connects to the charge e.

4 Gyromagnetic ratio g

For (8) and (9), the g-factor is exactly $g = 2$ [7, p.667]. This gyromagnetic ratio represents the magnetic to mechanical moment-ratio (13) for the Dirac equation without radiative corrections.

In general (radiative corrections or not), the intrinsic magnetic moment $\vec{\mu}$ is related to the spin vector $\vec{s} = \vec{S}/2$ through the equations [6, p.81]

$$\vec{\mu} = g\mu_B\vec{s} \quad \rightarrow \quad g\mu_B = \frac{\mu}{s} \tag{13}$$

where g is the g-factor and μ_B is the Bohr magneton

$$\mu_B = \frac{e\hbar}{2m_e c} = \frac{ec\hbar}{2m_e c^2} = \frac{ee_*^2}{2m_e c^2} = \frac{er_e}{2} \tag{14}$$

where r_e is the electron Compton radius. Although the g-factor in (13) is exactly 2 for the Dirac equation, there is an anomalous-moment increase to this value due to radiative corrections [6, p.298].

Note that for the Dirac particles where $g = 2$, (13) yields

$$\vec{\mu} = er_e\vec{s} \quad \rightarrow \quad \frac{\mu}{s} = er_e\,. \tag{15}$$

However, this is an unacceptable result for the Dirac proton; so (13) is replaced here by

$$\vec{\mu} = g\mu_c\vec{s} \quad \rightarrow \quad \frac{\mu}{s} = g\mu_c \tag{16}$$

where $\mu_c = er_e/2$ for the electron and $\mu_c = er_p/2$ for the proton. Thus the correct baseline moments, normalized by their common spin, for the Dirac particles are given by (16) with $g = 2$, where

$$\frac{\mu_e}{s} = er_e \quad \text{and} \quad \frac{\mu_p}{s} = er_p \tag{17}$$

are the electron and proton magnetic dipole moments.

5 Electron g-factor

When radiative corrections are included with (8) and (9), photon exchanges taking place within the vacuum state lead to a small increase in the electron g-factor and a large increase in the proton g-factor. Using $\alpha^{-1} = 137.0$ [7, p.722] for the inverse fine structure constant in the Schwinger calculation [8] [6, p.298], the relative change in the electron magnetic moment is

$$\frac{\delta\mu}{\mu} = \frac{g}{2} - 1 = \frac{e^2}{2\pi c\hbar} = \frac{1}{2\pi}\frac{e^2}{e_*^2}$$

$$= \frac{\alpha}{2\pi} = 0.001162 \tag{18}$$

where one of the e_*s in the squared coupling charge e_*^2 belongs to the electron and the other to any one of the Planck-particle cores within the degenerate PV state.

In the QED theory, the result in (18) is considered to be a first order (in $\alpha/2\pi$) [6, p.82] radiative correction. Like this first order correction, the higher-order corrections are difficult to calculate, but produce increasingly accurate results based on the QED methodology.

Using (18) to second order in $\alpha/2\pi$ leads to

$$\frac{g}{2} - 1 = \frac{\alpha}{2\pi} - \left(\frac{\alpha}{2\pi}\right)^2 = 0.001160 \tag{19}$$

where the experimental g-factor is [6, p.298]

$$\left(\frac{g}{2} - 1\right)_{exp} = 0.0011596 \approx 0.001160\,. \tag{20}$$

The fortuitous agreement between (19) and (20) depends upon the choice of α in the first paragraph.

6 Proton g-factor

The electron is thought to be a true point particle [6, p.82] because it contains no internal structure, as does the proton [9]. In the present context, however, it is appropriate to associate the "size" of the electron and proton with their Compton radii, where the corresponding proton structure constant is defined here by

$$m_p = \frac{r_e}{r_p}m_e \quad \rightarrow \quad \left(\frac{r_e}{r_p}\right) = \frac{m_p}{m_e} \approx 1836\,. \tag{21}$$

This suggests that the proton g-factor change be estimated from the electron change,

$$\frac{g}{2} - 1 = \left[\frac{\alpha}{2\pi} - \left(\frac{\alpha}{2\pi}\right)^2\right]\frac{r_e}{r_p} = 0.001160\,\frac{r_e}{r_p} = 2.13 \qquad (22)$$

where the experimental g-factor is [6, p.82]

$$\left(\frac{g}{2} - 1\right)_{exp} = 1.79\,. \qquad (23)$$

The agreement between (22) and (23) is remarkable, considering the large magnitude of r_e/r_p. It remains to be seen, however, whether or not (22) leads to something more substantial.

7 Summary and comments

It probably comes as a surprise that the charge associated with the Dirac equation and the Dirac particles is the coupling charge e_*, rather than the well known electron/proton charge e. That bewilderment is due to the collection of Planck particle cores that pervade the PV state. If there were no such pervasion, there would be no photon scattering taking place within the vacuum state and no resulting need for the coupling charge and the radiative corrections from the QED theory.

Sections 5 and 6 present calculations that suggest the PV theory may provide an aid to, or an alternative for, the difficult QED calculations that have been so spectacularly successful. That, of course, remains to be seen. But another hint that the PV theory may be a help is the Schwinger result in Section 5:

$$\frac{\alpha}{2\pi} = \frac{e^2/2\pi r_*}{m_* c^2} = \frac{e^2/2\pi r_p}{m_p c^2} = \frac{e^2/2\pi r_e}{m_e c^2} \qquad (24)$$

where, if r_* is the "radius" of the Planck-particle cores in the PV pervaded space, then $2\pi r_*$ is the "circumference" of the corresponding "spheres" surrounding those cores. Further work will be focused on developing a complete PV approach to the radiative correction phenomenon.

Feynman [10, p.129] notes that: "There is a most profound and beautiful question associated with the coupling constant, e—the amplitude for a real electron to emit or absorb a real photon. It is a simple number that has been experimentally determined to be close to -0.8542455. (My physicist friends won't recognize this number, because thy like to remember it as the inverse of its square: about 137.03597 with an uncertainty of about 2 in the last decimal place. It [the fine structure constant] has been a mystery ever since it was first discovered more than fifty years ago, and all good theoretical physicists put this number up on their wall and worry about it.)" The mystery of the fine structure constant α resides in the photon scattering that takes place within the pervaded PV state. It is also noted that the apparent electron emission/absorption of photons has its source in the pervaded nature of that state.

Appendix A: The γ and β matrices

The 4x4 γ, β, and α_i matrices used in the Dirac theory are defined here: where [6, p.91]

$$\gamma^0 \equiv \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \qquad (A1)$$

and ($j = 1, 2, 3$)

$$\gamma^j \equiv \beta\alpha_j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix} \qquad (A2)$$

and where I is the 2x2 unit matrix and

$$\alpha_j = \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix} \qquad (A3)$$

where the σ_j are the 2x2 Pauli spin matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (A4)$$

and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$. The zeros in (A1)–(A3) and (A5) are 2x2 null matrices.

The mc in (8) and (9) represents the 4x4 matrix

$$mc\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \qquad (A5)$$

and ψ is the 4x1 spinor matrix.

The zero on the right side of (9) represents the 4x4 null matrix and the zeros in (10) represent 2x2 null matrices. The S_j and σ_j in (10) are 2x2 matrices; so their parentheses represent 4x4 matrices.

The coordinates x^μ are

$$x^\mu = (x^0, x^1, x^2, x^3), \quad x^0 \equiv ct. \qquad (A6)$$

Received on February 28, 2019.

References

1. Davies P. Superforce: the Search for a Grand Unified Theory of Nature. Simon and Schuster, New York, 1984.
2. Daywitt W. C. A Model for Davies' Universal Superforce. *Galilean Electrodynamics*, 2006, v. 5, 83. See also www.planckvacuumDOTcom.
3. Daywitt W. C. The Trouble with the Equations of Modern Fundamental Physics. *American Journal of Modern Physics, Special Issue: Physics without Higgs and without Supersymmetry*, 2016, v. 5 (1-10, 22.
4. Carroll B. W., Ostlie D. A. An Introduction to Modern Astrophysics. Addison-Wesley, San Francisco, 2007.
5. Daywitt W. C. The Planck vacuum. Progress in Physics, 2009, v. 5, 20.
6. Gingrich D. M. Practical Quantum Electrodynamics. CRC, The Taylor & Francis Group, Boca Raton, 2006.
7. Leighton R. B. Principles of Modern Physics. McGraw-Hill, New York, 1959.
8. Schwinger J. On Quantum-Electrodynamics and the Magnetic Moment of the Electron. *Physical Review*, 1948, v. 73, 416.
9. Daywitt W. C. The Structured Proton and the Structureless Electron as Viewed in the Planck Vacuum Theory. *Progress in Physics*, 2015, v. 11 (2), 117.
10. Feynman R. P. QED: the Strange Theory of Light and Matter. Princeton University Press, Princeton, 1985.

A Derivation of Space and Time

Paul Bernard White

ORCID: 0000-0002-2681-3670. E-mail: pbwx@att.net

Four simple postulates are presented, from which we derive a (3+1)-dimensional structure, interpreted as ordinary space and time. We then derive further properties of space: isotropy and homogeneity; a rapid expansion within the first instant of time (*i.e.* inflation); and a continual and uniform expansionary pressure, due to a continual influx of (*non*-zero-point) energy that is uniformly distributed (*i.e.* dark energy). In addition, the time dimension is shown to have an "arrow". These results suggest that the four postulates may be fundamental to the construction of the physical universe.

1 Introduction

Systems that are based on information typically contain a basic information *element* and a basic information *structure*. In biological systems, for example, the basic information element is the nucleotide molecule, and the basic information structure is a *sequence* of nucleotides (*e.g.* a codon, or a gene). Likewise, for computer systems the basic information element is the bit, and the basic information structure is a sequence of bits (*e.g.* an 8-bit byte). And in natural language the basic information element is the letter or phoneme, and the basic information structure is a sequence of letters or phonemes (*e.g.* a word or a sentence).

Such systems must also have a way of translating or computing the information elements and structures into meaningful output. In biology this is accomplished by the operations of ribosomes, enzymes, *etc.*, acting on the nucleotide strings. For computers, the operations of logic gates on the bit strings typically perform this function. And in natural language the operations of lexical analysis, parsing, and context translate a string of letters/phonemes into meaning.

Similarly, if the *physical universe* is based on information (as many have speculated, *e.g.* [1–3]), then the following questions arise: (a) What is the basic information element for this system?; (b) what is the basic information structure for the system?; and (c) how are these elements and structures translated (or computed) into the meaningful output that we call the physical universe?

In answer to questions (a) and (b) above, I propose the following two postulates:

1. For creation of the physical universe, the basic information element is a type of *projection* – more specifically, *a projection from a prior level*.
2. The basic information structure is a *sequence* of such projections.

With respect to the first postulate, we may refer to both projections and levels as "elements" (or *basic* elements) of the system, but will reserve the term "basic *information* element" for the projections alone.

We now add two more postulates:

3. Each such projection is a *one*-dimensional *vector*, constituting a *different*, but related, one-dimensional space. (The basic relations between these projections/vectors are stated in the next postulate.)
4. Prior things (*e.g.* projections, levels, and constructions from them) are *independent* of subsequent things; and, conversely, subsequent things are *dependent* on prior things. (The terms prior, subsequent, dependent, and independent denote here *logical/ontological* relations. See *e.g.* [4].)

In [5], I use these four postulates (and two additional ones) to develop a model for the basic construction of the physical universe – including the construction of ordinary space and time themselves, the fundamental particles and interactions, *etc.* In the present paper, however, we will (for the sake of brevity) focus simply on constructing ordinary space and time, and their basic properties. That is, using the four postulates above, we will:

- derive a (3+1)-dimensional structure, interpreted as ordinary space and time
- show that the derived 3-dimensional space is isotropic and homogeneous, and that the time dimension has an "arrow"
- show that space undergoes a rapid expansion within the first instant of time (*i.e.* inflation)
- show that space undergoes a continual and uniform expansionary pressure, due to a continual influx of (*non*-zero-point) energy that is uniformly distributed (*i.e.* dark energy).

With respect to question (c) above, it will be shown that a method for translating sequences of projections into physical meaning is by taking into account the *relations* between projections – specifically, their dependence and independence relations (*i.e.* postulate 4). Once obtained, the above (bulleted) results can then be said to support the proposition that *the four stated postulates are fundamental to the construction of the physical universe*.

From now on, we will often refer to the model for constructing the ***p***hysical universe, developed herein, as system P.

2 Levels, projections, and relations: the structure and basic properties of system P

To construct our model for the physical universe (*i.e.* system P), we must begin with a *state* at which the things of the universe do not exist (otherwise our construction would be circular), *i.e.* a state that is absent the energy, elementary particles, and even space and time, as we know them. We will call this state *level 0 of system P*, or just *level 0*. We do not, however, presume that level 0 is a state of nothingness, or that nothing exists at level 0. We merely claim that nothing that comes into being with the construction of the physical universe exists at level 0; for level 0 is by definition a state that is immediately *prior* to the construction of the physical universe.

Recalling our first three postulates, we say that a projection from level 0, to be denoted as $\mathbf{p}_0$, generates a *new* state, which we call *level 1*. Likewise, a projection from level 1, denoted as $\mathbf{p}_1$, generates another new state, which we call *level 2*. And a projection from level 2, denoted as $\mathbf{p}_2$, yields *level 3*; and so on. So, in general, the projection $\mathbf{p}_k$ represents a sort of *displacement* from level k that generates level $k+1$ (for $k = 0, 1, 2, \ldots$); thus, relative to each other, level k is *prior*, and level $k+1$ is *subsequent*; also, relative to each other, $\mathbf{p}_k$ is prior, and $\mathbf{p}_{k+1}$ is subsequent. (Again, the terms "prior" and "subsequent" refer to logical/ontological priority and subsequence.)

In Fig. 1, where levels are represented by horizontal lines, and projections are represented by vertical arrows from a prior level to the next subsequent level, we illustrate the construction of levels 1 through 3 via the projections $\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_2$. To the right of each level in Fig. 1 is shown the sequence of projections that is required to construct that level (the round brackets indicate a sequence, as is common in mathematics). Thus, the sequences of projections that are required to create levels 0, 1, 2, and 3 are ( ), $(\mathbf{p}_0)$, $(\mathbf{p}_0, \mathbf{p}_1)$, and $(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2)$, respectively; moreover, the latter sequence constructs *all* of the levels (above level 0) in Fig. 1.

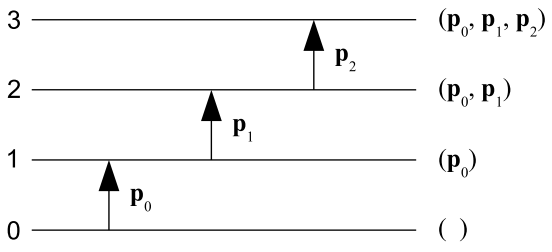

Fig. 1: Construction of levels 1 through 3 of system P via the projection sequence $(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2)$. The projection sequence that is required to construct a given level is shown to the right of that level.

As just described, the *order* of construction in system P starts with level 0 at the bottom of Fig. 1 and proceeds in the upward direction. Thus, level 0 is *prior* to all other elements (levels or projections) in system P, and *subsequent* to none; $\mathbf{p}_0$ is subsequent to level 0, but prior to level 1, $\mathbf{p}_1$, level 2, *etc.*; and so on. So, in general, a given element x in system P is subsequent to everything below it in Fig. 1, but prior to everything above it. By postulate 4, this means that element x is *dependent* on everything below it in Fig. 1, but *independent* of everything above it. Thus, for example, level 0 is independent of all other elements in system P, and dependent on none.

Since level 0 is our *starting point* (or starting *state*) for constructing system P, then we must say that it is a *nonconstructed* element of that system, whereas the subsequent projections and levels ($\mathbf{p}_0$, level 1, $\mathbf{p}_1$, level 2, *etc.*) are *constructed* elements of system P. So anything subsequent to level 0 is a *constructed* entity of the system.

2.1 Some properties of system P

Let x be a thing of system P (*e.g.* x is a level, a set of one or more projections, or something constructed from them). By postulate 4, things that are subsequent to x are (logically/ontologically) *dependent* on x. Such dependence implies that x is **in effect**, effective, operative, or **operant** at those subsequent things; or, alternatively, we say that those subsequent/dependent things are *within the* **scope** *of* x. Conversely, since things that are *prior* to x are *independent* of it, we say that x is *not* in effect or operant at those prior things; or, alternatively, we say that those prior/independent things are *not* within the scope of x. All of this is summarized in what will be called the **scope rule** for system P, stated as follows:

> A given thing in system P is in effect/operant at (*i.e.* contains within its scope) those things which are *subsequent*, and is not in effect at (does not contain within its scope) those things which are *prior*.

From this we may deduce the following corollary to the scope rule:

> A given *element* in system P (*i.e.* a projection or level) is in effect/operant at (contains within its scope) those elements that are *above* it in Fig. 1, and is not in effect at (does not contain within its scope) those elements that are *below* it in Fig. 1.

Thus, for example, since all of the *constructed* elements of system P (*i.e.* $\mathbf{p}_0$, level 1, $\mathbf{p}_1$, level 2, *etc.*) are subsequent to level 0 (or, conversely, level 0 is prior to them), then level 0 is in effect/operant at all of those things; or, all of those things are within the scope of level 0. Likewise, $\mathbf{p}_1$, level 2, $\mathbf{p}_2$, and level 3 are within the scope of level 1; but level 0 is *not* within the scope of level 1. And so on.

Since $\mathbf{p}_k$ is *not* in effect at level k, but *is* in effect at level $k+1$, then level $k+1$ represents the *state* at which the projection $\mathbf{p}_k$ first comes into effect; by the scope rule, $\mathbf{p}_k$ then *stays* in effect for all subsequent levels. Thus, the projection $\mathbf{p}_0$ first comes into effect at level 1, and stays in effect for levels 2 and 3; likewise, $\mathbf{p}_1$ first comes into effect at level 2, and stays in

effect for level 3. Let us say that the level at which a projection first comes into effect is its *native level*. Thus, level 1 is the native level for $\mathbf{p}_0$; level 2 is the native level for $\mathbf{p}_1$; and so on. That is, the native level for $\mathbf{p}_k$ is level $k+1$. Moreover, the concept of native level can be extended to things that are *constructed from* projections; thus, for example, something that is constructed using $\mathbf{p}_0$ and $\mathbf{p}_1$ (and no other projections) is native to level 2, since those two projections are first *jointly* in effect at that level. We note also that the projections that are in effect/operant at a given level are the *same* as the ones that are required to *construct* that level (as described earlier, and as listed in the sequences to the right of each level in Fig. 1).

In constructing the sequence of projections $(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2)$, since any projections that are in effect at level k are also in effect at the subsequent level $k+1$, then we can think of the latter level as *inheriting* all of the projections that are in effect at the former level. And since this is true of projections, then it is also true of anything that is associated with or constructed from them. This aspect of system P – whereby that which is in effect at one level (or, if you will, *generation*) is passed on to the next subsequent level (and thus, by extension, to *all* subsequent levels) – will be called the **inheritance rule**.

3 Constructing space and time in system P

Following postulate 3, let us model each projection as a one-dimensional *vector*; *i.e.* we model each $\mathbf{p}_k$ $(k = 0, 1, 2)$ as a one-dimensional vector going from level k to level $k+1$. Thus, $\mathbf{p}_0$ is a one-dimensional vector from level 0 to level 1; $\mathbf{p}_1$ is a one-dimensional vector from level 1 to level 2; and so on. These vectors are represented graphically by the vertical arrows in Fig. 1.

Moreover, each $\mathbf{p}_k$ constitutes a *different* one-dimensional space. Though they are different in this respect, the $\mathbf{p}_k$ are nevertheless *related* by the dependence and independence relations that have been postulated and discussed.

3.1 Constructing a (3+1)-dimensional structure at level 2 (and above)

Since $\mathbf{p}_0$ is the only projection in effect at level 1, and since (by postulate 3) it is *one* dimensional, then it is fair to say that system P is *one dimensional* at level 1.

Since both $\mathbf{p}_0$ and $\mathbf{p}_1$ are in effect at level 2, and since (by postulate 3) each of these constitutes a different one-dimensional space, then it might seem – at first glance – that system P should be *two* dimensional at level 2. But this would be wrong.

To get the correct dimensionality at level 2, we must take into account the *relations* between $\mathbf{p}_0$ and $\mathbf{p}_1$, as per postulate 4 – *i.e.* the fact that $\mathbf{p}_0$ is *independent* of $\mathbf{p}_1$, and that this relation is *asymmetric* ($\mathbf{p}_1$ is *dependent* on $\mathbf{p}_0$). Since $\mathbf{p}_0$ and $\mathbf{p}_1$ are *vectors*, we interpret that these relations imply a kind of (asymmetric) *linear* independence, with the following property: from the perspective of $\mathbf{p}_1$, the vector $\mathbf{p}_0$ may be collinear with $\mathbf{p}_1$, but is also free to be *noncollinear* with $\mathbf{p}_1$. With these considerations in mind, we ask the question: What is the *direction* of $\mathbf{p}_0$ with respect to $\mathbf{p}_1$? Or, in other words, how does $\mathbf{p}_0$ "look" relative to $\mathbf{p}_1$?

Since $\mathbf{p}_0$ may be both collinear and noncollinear with $\mathbf{p}_1$ (from the latter's perspective), then $\mathbf{p}_0$ may have a component parallel to $\mathbf{p}_1$, and may also have a component *perpendicular/orthogonal* (*i.e.* at 90 degrees) to $\mathbf{p}_1$. But, by symmetry, the perpendicular component can be anywhere in a two-dimensional *plane* orthogonal to $\mathbf{p}_1$. The two dimensions of this orthogonal plane, plus the one dimension parallel to $\mathbf{p}_1$, makes three dimensions. Thus, from the viewpoint of $\mathbf{p}_1$ (and from the perspective of level 2), $\mathbf{p}_0$ has *three* dimensions; *i.e.* $\mathbf{p}_0$ constitutes a *three*-dimensional space (whereas, recall that $\mathbf{p}_0$ has only one dimension at level 1). We might say, therefore, that the view of $\mathbf{p}_0$ from the perspective of $\mathbf{p}_1$ "bootstraps" the former from a one-dimensional vector into a three-dimensional space.

In summary, to construct its interpretation of $\mathbf{p}_0$, we can think of $\mathbf{p}_1$ as applying postulates 3 and 4 in succession: first, by postulate 3, $\mathbf{p}_0$ is a one-dimensional vector; second, by postulate 4, $\mathbf{p}_0$ is independent of $\mathbf{p}_1$ – which allows the former to have a component that is orthogonal to $\mathbf{p}_1$, with the result that $\mathbf{p}_1$ sees $\mathbf{p}_0$ as three dimensional.

Conversely, we can ask, how does $\mathbf{p}_1$ "look" relative to $\mathbf{p}_0$? Since $\mathbf{p}_1$ is *dependent* on $\mathbf{p}_0$, then the former is *not* free to have a component that is orthogonal to the latter, and so $\mathbf{p}_0$ sees $\mathbf{p}_1$ as being collinear; or, more simply, $\mathbf{p}_0$ sees $\mathbf{p}_1$ strictly as per postulate 3: as a *one*-dimensional vector.

So, at level 2 we have the three dimensions of $\mathbf{p}_0$, plus the one dimension of $\mathbf{p}_1$, for a total of *four* dimensions. Since system P is a model for constructing the physical universe, we interpret that the three dimensions of $\mathbf{p}_0$ are just the three dimensions of *ordinary space*, and the one dimension of $\mathbf{p}_1$ is the dimension of *time*; thereby yielding at level 2 the signature 3+1 space and time dimensions of our experience. The dimension of time, therefore, being a consequence of $\mathbf{p}_1$ (and $\mathbf{p}_0$), does not exist at levels 0 and 1, but only comes into existence at level 2; likewise, since ordinary, three-dimensional space is a consequence of $\mathbf{p}_0$ and $\mathbf{p}_1$, it also does not exist at levels 0 and 1, but only comes into existence at level 2.

Note that, although $\mathbf{p}_0$ itself is independent of $\mathbf{p}_1$, the triple dimensionality of $\mathbf{p}_0$ at level 2 is *not* independent of $\mathbf{p}_1$. That is, in the process described above, $\mathbf{p}_0$ only manifests as *three* dimensional when it is related to, or juxtaposed with, $\mathbf{p}_1$. Thus, the triple dimensionality of $\mathbf{p}_0$ at level 2 (*i.e.* the triple dimensionality of ordinary space) is in fact *dependent* on $\mathbf{p}_1$. Conversely, both $\mathbf{p}_0$ and $\mathbf{p}_1$ are *prior* to, and thus independent of, ordinary space.

We have shown, among other things, that $\mathbf{p}_0$ manifests differently at levels 1 and 2. At level 1 it is *one* dimensional. But when juxtaposed with $\mathbf{p}_1$ at level 2 it manifests as a *three*-dimensional space. Note that $\mathbf{p}_0$ itself does not change from level to level: it represents a projection from level 0 to level 1

wherever it appears (*i.e.* wherever it is in effect). This is analogous to *e.g.* the G nucleotide in biology, which is always the same molecule wherever it appears, but yields a different output (*i.e.* amino acid) depending on what other nucleotides/letters it is juxtaposed with in a sequence. In other words, like the letter G in a DNA sequence, the *meaning* of $\mathbf{p}_0$ is context dependent; which is just what we might expect for an element of a *language*, thus supporting our earlier notion that the basis of the physical universe is, to some degree at least, informational in nature.

We might say that level 2 has *two* dimensions as *input* (one dimension for $\mathbf{p}_0$, plus one for $\mathbf{p}_1$), but has *four* dimensions as *output* – three for $\mathbf{p}_0$, and one for $\mathbf{p}_1$. Which brings us back to question (c) in the introduction: How are the basic information elements of the model (which at level 2 are the inputs $\mathbf{p}_0$ and $\mathbf{p}_1$) translated (or, if you will, computed) into the meaningful output that we call the physical universe? We now see that at least a partial answer is that the *relations* between prior and subsequent elements are what translate them into meaningful output. In the present case, the independence relation between $\mathbf{p}_0$ and $\mathbf{p}_1$ at level 2 translates/transforms the manifestation of the former from a one-dimensional entity into a three-dimensional space.

We can thus say that the construction of each space at level 2 requires the participation of an *observer*, in the sense that $\mathbf{p}_1$ "observing" $\mathbf{p}_0$ constructs ordinary, three-dimensional space, and $\mathbf{p}_0$ "observing" $\mathbf{p}_1$ constructs one-dimensional time. With ordinary space *itself* constructed by an observation of sorts, it becomes more plausible that *e.g.* the *position* of an object *within* ordinary space might also be constructed by some type of observation, as seems to be the case in quantum mechanics (more about that in [5]).

The projections $\mathbf{p}_0$ and $\mathbf{p}_1$ are also operant at *level 3* (as per the scope rule), and the relations between them are the same as at level 2 (*i.e.* $\mathbf{p}_0$ is independent of $\mathbf{p}_1$, but not the converse). Thus, at level 3 – as at level 2 – $\mathbf{p}_0$ will appear to $\mathbf{p}_1$ as a three-dimensional space (*i.e.* ordinary space), and $\mathbf{p}_1$ will appear to $\mathbf{p}_0$ as a one-dimensional space (*i.e.* time). In other words, the spaces that exist at level 2 also exist at level 3. Indeed, as per the inheritance rule, we might say that level 3 *inherits* these spaces from level 2; or, more precisely, level 3 inherits $\mathbf{p}_0$, $\mathbf{p}_1$, and the relations between them from level 2, and uses them to *construct* ordinary space and time.

3.2 Isotropy and homogeneity of space

Recall that ordinary, three-dimensional space is created when $\mathbf{p}_0$ is viewed from the perspective of $\mathbf{p}_1$. So it follows that (a) the creation/construction of ordinary space is *dependent* on $\mathbf{p}_0$ and $\mathbf{p}_1$; and (b) $\mathbf{p}_0$ and $\mathbf{p}_1$ are *prior to*, and thus (by postulate 4) *independent of*, ordinary space.

Suppose now that an outcome of constructing ordinary space is that $\mathbf{p}_0$ (or $\mathbf{p}_1$) manifests with a particular orientation or direction within that space. Since this would make $\mathbf{p}_0$ (or $\mathbf{p}_1$) functionally dependent on ordinary space, and thus contradict (b) above, we conclude that the construction of ordinary space cannot result in $\mathbf{p}_0$ (or $\mathbf{p}_1$) having a particular direction/orientation within that space. Presumably, then, there is no way for the process that constructs ordinary space to establish a distinctive (*i.e.* special or preferred) direction within that space. We thus conclude that, as constructed above, ordinary space is perfectly *isotropic*.

Now suppose that an outcome of constructing ordinary space is that $\mathbf{p}_0$ (or $\mathbf{p}_1$) manifests with a particular *position* within that space. This, again, would make $\mathbf{p}_0$ (or $\mathbf{p}_1$) functionally dependent on ordinary space and thereby contradict (b) above; and so we conclude that the construction of ordinary space cannot result in $\mathbf{p}_0$ (or $\mathbf{p}_1$) having a particular position within that space. Presumably, then, the process that constructs ordinary space cannot establish a distinctive (*i.e.* special or preferred) position within that space. We thus conclude that, as constructed above, ordinary space is perfectly *homogeneous*.

In addition, the construction of ordinary space cannot result in either $\mathbf{p}_0$ or $\mathbf{p}_1$ manifesting as *vectors*, or *vector fields*, within that space; for if they did, then these projections/vectors would be functionally dependent on ordinary space, which would again contradict (b). Given that *vector* fields have been ruled out, it seems we have little choice but to assume that $\mathbf{p}_0$ and $\mathbf{p}_1$ manifest within ordinary space as uniform *scalar* fields – *uniform*, because any *non*uniformity would make the manifestations of $\mathbf{p}_0$ or $\mathbf{p}_1$ functionally dependent on ordinary space, which would, again, violate/contradict their independence from that space. Presumably, the uniform scalar field for $\mathbf{p}_0$ is just (raw, unstructured) ordinary space itself, and the uniform (one-dimensional) scalar field for $\mathbf{p}_1$ is just proper time.

Lastly, let us recall that $\mathbf{p}_0$ sees $\mathbf{p}_1$ as a one-dimensional *vector*. This, presumably, would impart some *directionality* to $\mathbf{p}_1$ – which, as we have concluded, could not manifest as a direction within ordinary space. Since $\mathbf{p}_1$ has been associated with *time*, we interpret that this directionality of $\mathbf{p}_1$ (with respect to $\mathbf{p}_0$) is just the "arrow" of time.

3.3 Rapid expansion of space within the first instant of time

Recall that $\mathbf{p}_0$ at level 1 is *one* dimensional – having, let us say, a length of p_0. The time dimension, being a result of $\mathbf{p}_1$, does not exist at this level/stage. Given that a one-dimensional object has *zero* volume, then the physical universe at this stage of development has a volume of zero.

Since the time dimension comes into existence with the projection $\mathbf{p}_1$, then the *advent* of $\mathbf{p}_1$ defines the time point $t = 0$, at which point p_0 has the value $p_0(t = 0)$, which may be denoted as $p_{0,0}$. So, at exactly $t = 0$, or within the first instant after it, the existence/perspective of $\mathbf{p}_1$ causes $\mathbf{p}_0$ to manifest as *three*-dimensional ordinary space, with a volume on the

order of $p_{0,0}^3$. Thus the volume of ordinary space goes from zero to around $p_{0,0}^3$ within a time interval of zero, or near-zero, length – which constitutes a potentially very large, perhaps infinite, rate of spatial expansion. I propose, therefore, that this rapid spatial expansion, triggered by the advent of $\mathbf{p}_1$ at $t = 0$, is the process known as *inflation* [6].

Note that, under the above mechanism, inflation has a natural beginning: the advent of $\mathbf{p}_1$ at $t = 0$. And it also has a natural *ending*: it ends when the volume of ordinary space is around $p_{0,0}^3$. So inflation only lasts for the time (if any) that it takes (from the perspective of $\mathbf{p}_1$) for the *one*-dimensional space of length $p_{0,0}$ to become the *three*-dimensional space of approximate volume $p_{0,0}^3$.

3.4 A continual influx of energy associated with $\mathbf{p}_0$, yielding a continual and uniform expansionary pressure on space

In constructing the sequence $(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2)$ for system P, let us assume that *energy* is needed to create each of the projections $\mathbf{p}_k$ (for $k = 0, 1, 2$). We can think of this energy as being stored along the length of $\mathbf{p}_k$, and/or as being stored in the *level* that is created by $\mathbf{p}_k$. So we can speak of "$\mathbf{p}_k$ energy", and/or we can speak of the energy, E_{k+1}, that $\mathbf{p}_k$ inputs into level $k + 1$. Thus, $\mathbf{p}_0$ is a *process* through which energy E_1 is input into level 1 of system P. Likewise, $\mathbf{p}_1$ is a process that inputs energy E_2 into level 2; and $\mathbf{p}_2$ is a process that inputs energy E_3 into level 3. The total energy, E_t, that is input into system P is therefore $E_t = E_1 + E_2 + E_3$. We assume that all of these energies are nonzero and positive, so the energy of system P at level 1 and above, due to contributions from the sources mentioned, is positive.

Now recall that the dimension of time is associated with $\mathbf{p}_1$. Since $\mathbf{p}_1$ does not exist at levels 0 and 1, then time also does not exist there; *i.e.* all time intervals are zero at those levels. Indeed, we can say that levels 0 and 1 are *independent of time*. But $\mathbf{p}_1$ *does* exist at level 2 and above; so time exists there, and *all* time intervals at those levels are *nonzero* (and presumably positive).

Thus, at level 1, energy is nonzero, but time is zero. At level 2 (and above), however, both energy and time (intervals) are *nonzero*. Consequently, at level 2 and above, the *product* of energy and time – the quantity known as *action* – is nonzero, and thus has a positive lower bound; *i.e.* at level 2 (and above) the action is *quantized*. We thus have the derivation of an action *quantum*, which we interpret to be the basis for the empirically-known "quantum of action", commonly referred to as *Planck's constant*, and denoted as h.

In the present model, therefore, the quantum of action, h, depends on both $\mathbf{p}_0$ and $\mathbf{p}_1$, and so does not exist at levels 0 and 1, but only comes into being at level 2. Thus, quantum mechanics, which is based on h, also comes into being at level 2 of system P. And therefore, due to the scope rule, both h and quantum mechanics are operant at level 2 and above; *i.e.* they are *native* to level 2.

The presence of h at levels 2 and 3 can, and we assume *does*, partition the energies E_2 and E_3 into a multiplicity of smaller chunks, yielding *many* objects/particles at those levels. The *absence* of h at level 1, however, means that the energy E_1 *cannot* be broken into chunks; and so the energy E_1 at level 1 constitutes a *single*, continuous entity. In addition, given that time exists at levels 2 and 3, we assume (as per special relativity) that the particles at those levels possess *mass*; and, given that time does *not* exist at level 1, we assume that the single entity at level 1 is *massless*. Furthermore, in [5] it is shown that the objects at level 3 have *internal* structure, whereas the objects at level 2 are *structureless*. These results lead us to identify the level-3 objects as *baryons*, and the level-2 objects as *leptons*. Moreover, since time exists at levels 2 and 3, then the input of energies (E_2 and E_3) into those levels can be, and we assume *is*, time limited – yielding a *finite* number of baryons at level 3, and a finite number of leptons at level 2.

Recall now that $\mathbf{p}_0$ is native to level 1, but *time* is native to *level 2*. Thus, $\mathbf{p}_0$ is *prior* to time. By postulate 4, this means that the $\mathbf{p}_0$ process, which pumps energy E_1 into level 1, is *independent of time*, and is therefore a *continual* process – *i.e.* it never stops, and so it must be happening right now. Consequently, the quantity E_1 is *always* increasing. Moreover, since E_1 is the energy of $\mathbf{p}_0$ at level 1, and since $\mathbf{p}_0$ (as seen by $\mathbf{p}_1$) *is* ordinary space, then it is clear that E_1 is just the energy of space itself. Hence, an always-increasing E_1 should yield a continual *expansionary* pressure on space. Indeed, an increase in E_1 may produce an increase in the *length* of $\mathbf{p}_0$, and thus an increase in p_0^3 (the size/volume of the physical universe).

Suppose now that the $\mathbf{p}_0$ process distributes its energy E_1 *nonuniformly* within space. This would make that process (and thus $\mathbf{p}_0$ itself) functionally *dependent* on space, and thereby contradict statement (b) in section 3.2. Consequently, the energy E_1 must be distributed *uniformly* throughout space. Since this process is also independent of time, then it is *constant in time*. So the continual influx of E_1 energy into the system via the $\mathbf{p}_0$ process yields an input of energy per unit volume of space that is uniform throughout space, and constant in time; in other words, E_1 yields a *cosmological constant*.

Taken all together, the above results suggest that we interpret E_1 to be the phenomenon known as *dark energy* [7]; *i.e.*

$$\text{dark energy} = E_1.$$

Moreover, since the $\mathbf{p}_0$ process and E_1 are *level-1* phenomena, but h only becomes operant at *level 2*, then dark energy/E_1 is prior to – and thus independent of – h and quantum mechanics, and so is *not* a zero-point energy.

4 Conclusion

A truly fundamental model of the universe must *derive* space and time – not just take them as given. Firstly, such a model

should derive the (3+1)-dimensionality of space and time, and the isotropy and homogeneity of space. Secondly, since inflation and dark energy are likely to be important factors in the construction of space, then the model should also derive them. As shown above, the present model meets these basic criteria, which indicates that the four stated postulates may be fundamental to the construction of the physical universe.

Acknowledgments

I would like to acknowledge the following people for enabling the research for this paper to take place: My late mother and late father, Margean A. White and Phillip B. White; my brother, Jeffrey C. White; John P. A. Higgins; and the late Louise L. Hay (The Hay Foundation).

Received on March 6, 2019

References

1. Wheeler J. A. It From Bit. In his: At Home in the Universe. AIP Press, 1994.
2. Wheeler J. A. The Computer and the Universe. *Int. J. Theor. Phys.*, 1982, v. 21, 6–7.
3. Davies P., Gregersen N. H., eds. Information and the Nature of Reality. Cambridge University Press, Cambridge, 2010.
4. Van Cleve J. Dependence. In: Audi R., ed. The Cambridge Dictionary of Philosophy. Cambridge University Press, Cambridge, 1995, p. 191.
5. White P. B. A Model for Creation: Part I. 2019, https:www.academia.edu/38496134.
6. Guth A. The Inflationary Universe. Perseus Books, Reading, MA, 1997.
7. Carroll S. M. Dark Energy and the Preposterous Universe. arXiv:astro-ph/0107571v2.

Twin Universes: a New Approach

Patrick Marquet
E-mail: patrick.marquet6@wanadoo.fr

In this article, we derive a differential form of Einstein's field equations using Cartan's free coordinates calculus. Under this form, we see that it is possible to infer another set of field equations dual to the original one and which displays a negative sign. We may then relate this system to the equations sustaining the twin Universe of the Janus Cosmological Model developed by the astrophysicist J.-P. Petit.

Introduction

As early as 2014, the astrophysicist J.-P. Petit put forward a model of Universe which harbors two fields equations with two sources: it is referred to as *The Janus Cosmological Model* (JCM) [1] which is inspired by the twin Universes theory first proposed by A. Sakharov [2].

Such a bi-metric is shown to account for the Dark Energy description and other unsolved observational data [3], provided one distinguishes our Universe as filled with positives masses and energies, from another wherein negative masses and negative energies are assigned to.

From the quantum physics perspective, negative energies have always played an unsavory role.

However, following a recent publication, it appears that both negative energies and masses are physically compatible if the time reversal operator is kept unitary within the Dirac formalism [4].

This considerable mathematical progress lends support to the Janus Model which relies on this symmetry.

So far, the few theories exhibiting two opposite metrics have been arbitrarily assumed as a "natural" hypothesis with the confidence that subsequent results would eventually corroborate this postulate. In this paper, we tackle the problem at the very early stage: With the aid of the Cartan calculus and using the Hodge star operation, we rewrite the Einstein's field equations under a differential form.

With this preparation, we naturally infer another set of field equations which displays a negative sign. This differential procedure thus provides a straightforward basis wherefrom the *Janus Model* can be substantiated.

Notations

Space-time: Greek indices α, β run from 0, 1, 2, 3. Space-time signature: -2. In the present text, $\varkappa$ is the Einstein's constant: $8\pi G/c^4$ where G is Newton's gravitational constant, although we adopt here $c = 1$.

1 Differential form of Einstein's field equations

1.1 The Cartan procedure

Let us consider a 4-pseudo-Riemannian manifold referred to a general basis e_α. The dual basis θ^β of one-forms are related to the local (Roman) coordinates $\{a\}$ by:

$$\theta^\beta \;=\; a^\beta_a \,\mathrm{d}x^a. \tag{1.1}$$

The (a^β_a) are called *vierbein* or *tetrad fields* [5].

We next define the *Cartan procedure*, a powerful coordinates free calculus which is extensively used in the foregoing.

Let us define the *connection forms* by:

$$\Gamma^\alpha_\beta \;=\; \left\{ {\alpha \atop \gamma\,\beta} \right\} \theta^\gamma. \tag{1.2}$$

The first Cartan structure equation is related to the torsion by [6, p.40]:

$$\Omega^\alpha \;=\; \frac{1}{2}\, T^\alpha_{\ \gamma\,\delta}\, \theta^\gamma \wedge \theta^\delta \;=\; \mathrm{d}\theta^\alpha \;+\; \Gamma^\alpha_\gamma \wedge \theta^\gamma, \tag{1.3}$$

where $T^\alpha_{\ \gamma\,\delta} \;=\; \frac{1}{2}\big[\Gamma^\alpha_{\ [\gamma\,\delta]} - \Gamma^\alpha_{\ [\delta\,\gamma]}\big]$ is the torsion tensor.

In the Riemannian framework alone, it reduces obviously to:

$$\mathrm{d}\theta^\alpha \;=\; -\Gamma^\alpha_\gamma \wedge \theta^\gamma. \tag{1.4}$$

The *second Cartan structure equation* is defined as [6, p.42]:

$$\Omega^\alpha_\beta = \frac{1}{2}\, R^\alpha_{\beta\gamma\delta}\, \theta^\gamma \wedge \theta^\delta = \mathrm{d}\Gamma^\alpha_\beta + \Gamma^\alpha_\gamma \wedge \Gamma^\gamma_\beta, \tag{1.5}$$

$R^\alpha_{\beta\gamma\delta}$ are here the curvature tensor components.

Defining the absolute exterior differential D of a tensor valued p-form of type (r, s)

$$(D\phi)_{j_1\ldots j_s}{}^{i_1\ldots i_r} = \mathrm{d}\phi_{j_1\ldots j_s}{}^{i_1\ldots i_r} + \Gamma^{i_1}_k \wedge \phi_{j_1\ldots j_s}{}^{k\,i_2\ldots i_r} + \ldots - \Gamma^{\ k}_{j_1} \wedge \phi_{k\,j_2\ldots j_s}{}^{i_1\ldots i_r} - \ldots$$

we can write for example the *Bianchi identities* in a very simple way as:

$$D\Omega^\alpha \;= \Omega^\alpha_\beta \wedge \theta^\beta, \tag{1.6}$$

$$D\Omega^\alpha_\beta \;= 0. \tag{1.7}$$

1.2 The Einstein equations

1.2.1 The Einstein action

We first recall the *Hodge star* operator definition for an oriented n-dimensional *pseudo-Riemannian manifold* $(\mathrm{M}, \mathbf{g})$ whose volume element determined by $\mathbf{g}$ is:

$$\boldsymbol{\eta} \;=\; \sqrt{-\mathbf{g}}\; \theta^0 \wedge \theta^1 \wedge \theta^2 \wedge \theta^3.$$

Let $\Lambda_k(E)$ be the subspace of completely *antisymmetric multilinear forms* on the real vector space E.

The *Hodge star operator* * is a *linear isomorphism* *: $\Lambda_k(E) \to \Lambda_{n-k}(M)$ $(k \leqslant n)$. If $\theta^0, \theta^1, \theta^2, \theta^3$ is an oriented basis of 1-forms, this operator is defined by:

$$^*(\theta^{i_1} \wedge \theta^{i_2} \wedge \ldots \wedge \theta^{i_k}) = \\ = \frac{\sqrt{-g}}{(n-k)!} \left[\epsilon_{j_1 \ldots j_n} \, g^{j_1 i_1} \ldots g^{j_k i_k} \, \theta^{j_{k+1}} \wedge \ldots \wedge \theta^{j_n} \right]. \quad (1.8)$$

With this preparation, the Einstein action simply reads:

$$^*R \;=\; R\,\boldsymbol{\eta}\,. \quad (1.9)$$

We shall need this action expressed in terms of tetrads.

Proof: With $\sigma^{\mu\nu} = {}^*(\theta^\mu \wedge \theta^\nu)$ and taking into account (1.8) we have

$$\sigma_{\beta\gamma} \wedge \Omega^{\beta\gamma} \;=\; \frac{1}{2}\,\sigma_{\beta\gamma}\, R^{\beta\gamma}{}_{\mu\nu}\, \theta^\mu \wedge \theta^\nu$$

and

$$^*(\theta^\mu \wedge \theta^\nu) \;=\; \frac{1}{2}\,\boldsymbol{\eta}_{\beta\alpha\sigma\rho}\, g^{\beta\mu}\, g^{\alpha\nu}\, \theta^\sigma \wedge \theta^\rho$$

i.e.

$$\sigma_{\beta\gamma} \;=\; \frac{1}{2}\,\boldsymbol{\eta}_{\beta\gamma\sigma\rho}\, \theta^\sigma \wedge \theta^\rho. \quad (1.10)$$

Thus,

$$\sigma_{\beta\gamma} \wedge \theta^\mu \wedge \theta^\nu \;=\; \frac{1}{2}\,\boldsymbol{\eta}_{\beta\gamma\sigma\rho}\, \theta^\sigma \wedge \theta^\rho \wedge \theta^\mu \wedge \theta^\nu \;=\; (\delta^\mu_\beta\, \delta^\nu_\gamma \,-\, \delta^\mu_\gamma\, \delta^\nu_\beta)\boldsymbol{\eta}$$

and:

$$\sigma_{\beta\gamma} \wedge \Omega^{\beta\gamma} \;=\; \frac{1}{2}\,(\delta^\mu_\beta\, \delta^\nu_\gamma \,-\, \delta^\mu_\gamma\, \delta^\nu_\beta)\, R^{\beta\gamma}_{\mu\nu}\boldsymbol{\eta} \;=\; R\,\boldsymbol{\eta} \;=\; {}^*R.$$

Taking into account (1.10) let us now compute the absolute exterior differential:

$$D\sigma_{\beta\gamma} \;=\; \frac{1}{2}\,D\,(\boldsymbol{\eta}_{\beta\gamma\sigma\rho}\, \theta^\sigma \wedge \theta^\rho).$$

In an orthonormal system $\boldsymbol{\eta}_{\beta\gamma\sigma\rho}$ is constant and: $D\,\boldsymbol{\eta}_{\beta\gamma\sigma\rho} \;=\; 0$.

This reflects the fact that in the *Riemannian framework* (metric connection), orthonormality is preserved under parallel transport as well as the transported vector magnitude. Therefore:

$$D\sigma_{\beta\gamma} \;=\; \boldsymbol{\eta}_{\beta\gamma\sigma\rho} D\theta^\sigma \wedge \theta^\rho.$$

Now, bearing in mind that the basis θ^σ is a *tensor valued 1-form of type (1,0)*, the first structure equation reads [7]:

$$D\theta^\sigma \;=\; \Omega^\sigma$$

and

$$D\sigma_{\beta\gamma} \;=\; \boldsymbol{\eta}_{\beta\gamma\sigma\rho}\, \Omega^\sigma \wedge \theta^\rho = \Omega^\sigma \wedge \sigma_{\beta\gamma\sigma}\,.$$

The latter is zero for the torsion free Riemann connection: $D\sigma_{\beta\gamma} \;=\; 0$.

In the same way, we can show that

$$D\sigma^{\beta\gamma}{}_\alpha = \mathrm{d}\sigma^{\beta\gamma}{}_\alpha + \Gamma^\beta_{\ \delta} \wedge \sigma^{\delta\gamma}{}_\alpha + \Gamma^\gamma_{\ \delta} \wedge \sigma^{\beta\delta}{}_\alpha - \Gamma^\delta_{\ \alpha} \wedge \sigma^{\beta\gamma}{}_\delta \quad (1.11)$$

with

$$\sigma^{\beta\gamma}{}_\alpha \;=\; {}^*(\theta^\beta \wedge \theta^\gamma \wedge \theta_\delta),$$

(where all indices are raised or lowered with $g_{\alpha\beta}$ from $\boldsymbol{g} \;=\; g_{\alpha\beta}\,\theta^\alpha \otimes \theta^\beta$).

1.2.2 The Einstein field equations

From (1.10), we infer:

$$\sigma_{\beta\gamma\delta} \;=\; \boldsymbol{\eta}_{\beta\gamma\delta\lambda}\, \theta^\lambda. \quad (1.12)$$

Under the variation of $\delta\theta^\beta$ of the orthonormal tetrad fields, we have

$$\delta(\sigma_{\beta\gamma} \wedge \Omega^{\beta\gamma}) = \delta\sigma_{\beta\gamma} \wedge \Omega^{\beta\gamma} + \sigma_{\beta\gamma} \wedge \delta\Omega^{\beta\gamma}\,.$$

Now, using (1.10) and (1.12) yields:

$$\delta\sigma_{\beta\gamma} \;=\; \frac{1}{2}\,\delta(\boldsymbol{\eta}_{\beta\gamma\delta\lambda}\, \theta^\delta \wedge \theta^\lambda) \;=\; \delta\theta^\delta \wedge \sigma_{\beta\gamma\delta}\,.$$

Then, applying the varied second structure equation

$$\delta\Omega^{\beta\gamma} = \mathrm{d}\delta\Gamma^{\beta\gamma} + \delta\Gamma^\beta_{\ \eta} \wedge \Gamma^{\eta\gamma} + \Gamma^\beta_{\ \eta} \wedge \delta\Gamma^{\eta\gamma}$$

we obtain

$$\delta(\sigma_{\beta\gamma} \wedge \Omega^{\beta\gamma}) = \delta\theta^\gamma \wedge (\sigma_{\beta\gamma\delta} \wedge \Omega^{\beta\gamma}) + \mathrm{d}(\sigma_{\beta\gamma} \wedge \delta\Gamma^{\beta\gamma}) - \\ - \mathrm{d}\sigma_{\beta\gamma} \wedge \delta\Gamma^{\beta\gamma} + \sigma_{\beta\gamma} \wedge (\delta\Gamma^\beta_{\ \eta} \wedge \Gamma^{\eta\gamma} + \Gamma^\beta_{\ \eta} \wedge \delta\Gamma^{\eta\gamma}) \quad (1.13)$$

from the second line, we extract:

$$\mathrm{d}\sigma_{\beta\gamma} + \sigma_{\beta\gamma} \wedge (\Gamma^\eta{}_\gamma + \Gamma_{\beta\eta})$$

which is just: $D\sigma_{\beta\gamma}$. However, we know that: $D\sigma_{\beta\gamma} \;=\; 0$, and finally, the Einstein action variation is:

$$\delta(\sigma_{\beta\gamma} \wedge \Omega^{\beta\gamma}) \;=\; \delta\theta^\beta \wedge (\sigma_{\beta\gamma\delta} \wedge \Omega^{\gamma\delta}) \;+\; \mathrm{d}(\sigma_{\beta\gamma} \wedge \delta\Gamma^{\beta\gamma}) \quad (1.14)$$

(exact differential). The global Lagrangian density with matter is written:

$$\boldsymbol{L} \;=\; -\left(\frac{1}{2}\,\varkappa\right)^*\boldsymbol{R} \;+\; \boldsymbol{L}_{mat}\,.$$

Setting ${}^*T_\beta$ as the energy-momentum 3-form for *bare* matter we have the varied matter lagrangian density:

$$\boldsymbol{L}_{mat} \;=\; -\delta\theta^\beta \wedge {}^*T_\beta\,.$$

and taking into account (1.14) the global variation is:

$$\delta(\boldsymbol{L}) \;=\; -\delta\theta^\beta \wedge \left[\frac{1}{2}\,\varkappa\,\sigma_{\beta\gamma\delta} \wedge \Omega^{\gamma\delta} + {}^*T_\beta\right] \;+\; (\text{exact differential}).$$

We eventually arrive at the field equations under the differential form:

$$-\frac{1}{2}\sigma_{\beta\gamma\delta}\wedge\Omega^{\gamma\delta} \;=\; \varkappa\,{}^{*}T_{\beta}\,, \tag{1.15}$$

where T_α is related to the energy-momentum tensor $T_{\alpha\beta}$ by $T_\alpha = T_{\alpha\beta}\,\theta^\beta$.

In the same manner, one has: $G_\alpha = G_{\alpha\beta}\,\theta^\beta$ so that these identifications lead to the field equations with a source in the classical form:

$$G_{\alpha\beta} \;=\; R_{\alpha\beta} - \frac{1}{2}\,g_{\alpha\beta}\,R \;=\; \varkappa\,T_{\alpha\beta}\,, \tag{1.16}$$

$G_{\alpha\beta}$ is conserved but not $T_{\alpha\beta}$, therefore we should look for the appropriate r.h.s. tensor.

To this effect we start by reformulating (1.15) as

$$-\frac{1}{2}\,\Omega_{\beta\gamma}\wedge\sigma^{\beta\gamma}{}_{\alpha} \;=\; \varkappa\,{}^{*}T_{\alpha} \tag{1.17}$$

and we use the second structure equation under the following form

$$\Omega_{\beta\gamma} \;=\; \mathrm{d}\Gamma_{\beta\gamma} \;-\; \Gamma_{\mu\beta}\wedge\Gamma^{\mu}{}_{\gamma} \tag{1.18}$$

so as to obtain:

$$\mathrm{d}\Gamma_{\beta\gamma}\wedge\sigma^{\beta\gamma}{}_{\alpha} = \mathrm{d}\left(\Gamma_{\beta\gamma}\wedge\sigma^{\beta\gamma}{}_{\alpha}\right) + \Gamma_{\beta\gamma}\wedge d\sigma^{\beta\gamma}{}_{\alpha}\,. \tag{18bis}$$

Then using (1.11) in (1.18bis), we infer:

$$\begin{aligned}\mathrm{d}\Gamma_{\beta\gamma}\wedge\sigma^{\beta\gamma}{}_{\alpha} &= \mathrm{d}\left(\Gamma_{\beta\gamma}\wedge\sigma^{\beta\gamma}{}_{\alpha}\right) + \\ &+ \Gamma_{\beta\gamma}\wedge\left(\Gamma^{\beta}{}_{\delta}\wedge\sigma^{\delta\gamma}{}_{\alpha} - \Gamma^{\gamma}{}_{\delta}\wedge\sigma^{\beta\delta}{}_{\alpha} - \Gamma^{\delta}{}_{\alpha}\wedge\sigma^{\beta\gamma}{}_{\delta}\right).\end{aligned} \tag{1.19}$$

Adding the second contribution ($\Gamma^\alpha\ \gamma\wedge\Gamma^\gamma\ \beta$) of (1.18) to (1.19), we obtain the Einstein field equations in a new form:

$$-\frac{1}{2}\,\mathrm{d}(\Gamma_{\beta\gamma}\wedge\sigma^{\beta\gamma}{}_{\alpha}) \;=\; \varkappa\,({}^{*}T_{\alpha} \;+\; {}^{*}t_{\alpha}), \tag{1.20}$$

where

$${}^{*}t_{\alpha} \;=\; \left(-\frac{1}{2}\,\varkappa\right)\Gamma_{\beta\gamma}\wedge(\Gamma_{\delta\alpha}\wedge\sigma^{\beta\gamma\delta} - \Gamma^{\gamma}{}_{\delta}\wedge\sigma^{\beta\delta}{}_{\alpha}), \tag{1.21}$$

where ${}^{*}t_\alpha$ should be here interpreted as *energy* and *momentum 3-form of the gravitational field* generated by this matter.

Equation (1.20) readily implies the conservation law:

$$\mathrm{d}({}^{*}T_{\alpha} \;+\; {}^{*}t_{\alpha}) \;=\; 0\,. \tag{1.22}$$

Within the *Riemannian framework*, we know that the gravitational field cannot be localized, which is reflected by the fact that ${}^{*}t_\alpha$ does not transform as a tensor with respect to gauge transformations.

Indeed, as $\Gamma_{\beta\gamma}$ can be made zero at any given point of the Riemannian manifold, this 3-form vanishes.

To the 3-form ${}^{*}t_\alpha$ is thus associated the antisymmetric Einstein-Dirac pseudo-tensor $(\Theta^a_b)_{ED}$ [8].

In order to explicitly write down (1.20) with a *true* 3-form on the r.h.s., one should add the *3-form of the energy-momentum for the vacuum* denoted by $({}^{*}t_\alpha)_{vac}$.

Equation (1.22) eventually satisfies the conservation law:

$$\mathrm{d}[{}^{*}T_{\alpha} \;+\; ({}^{*}t_{\alpha})_{gravity}] \;=\; 0 \tag{1.23}$$

with:

$$({}^{*}t_{\alpha})_{gravity} \;=\; {}^{*}t_{\alpha} \;+\; ({}^{*}t_{\alpha})_{vac}\,. \tag{1.24}$$

To the 3-form $({}^{*}t_\alpha)_{vac}$ corresponds the tensor

$$(t_{\alpha\beta})_{vac} \;=\; \left(-\frac{1}{2}\,\varkappa\right)\Xi\,g_{\alpha\beta}\,, \tag{1.25}$$

where Ξ is the variable cosmological term which replaces the cosmological constant Λ as [9]:

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2}\,g_{\alpha\beta}\,R = \varkappa\left[T_{\alpha\beta} + (t_{\alpha\beta})_{ED}\right] + \Xi\,g_{\alpha\beta}\,. \tag{1.26}$$

2 Two opposite field equations

Since we deal with a Lorentzian manifold $n = 4$, repeated application of the duality operation *, gives:

$${}^{*}({}^{*}G_{\beta}) \;=\; -\,{}^{*}G_{\beta}\,, \tag{2.1}$$

$${}^{*}(\varkappa\,{}^{*}T_{\beta}) \;=\; -\,(\varkappa\,{}^{*}T_{\beta})\,. \tag{2.2}$$

The Cartan formalism thus allows for two "opposite" field equations to appear.

Can we find its physical meaning? A straightforward justification can be provided by the *Janus model* of J.P. Petit whose universes exhibit opposite energy/masses.

This model is characterized by two types of distinct metric tensors ${}^{(+)}g_{\mu\nu}$ and ${}^{(-)}g_{\mu\nu}$, which imply two distinct field equations:

$${}^{(+)}G_{\beta\mu} = {}^{(+)}R_{\beta\mu} - \frac{1}{2}\,{}^{(+)}g_{\beta\mu}{}^{(+)}R = \varkappa\left[{}^{(+)}T_{\beta\mu} + \varpi^{(-)}T_{\beta\mu}\right], \tag{2.3}$$

$${}^{(-)}G_{\beta\mu} = {}^{(-)}R_{\beta\mu} - \frac{1}{2}\,{}^{(-)}g_{\beta\mu}{}^{(-)}R = \varkappa\left[{}^{(-)}T_{\beta\mu} + \omega^{(+)}T_{\beta\mu}\right], \tag{2.4}$$

where ${}^{(+)}g_{\mu\nu}$ refers to positive mass/energy particles while ${}^{(-)}g_{\mu\nu}$ refers to negative mass/energy particles with the corresponding Ricci tensors ${}^{(+)}R_{\mu\nu}$ and ${}^{(-)}R_{\mu\nu}$.

Here ${}^{\pm}T_{\mu\nu}$ is the massive tensor which implicitly contains the gravitational field tensor defined from (1.24).

With our definition, we then have the obvious correspondences:

$${}^{*}G_{\beta} \;\to\; {}^{(+)}G_{\beta\mu}\,, \qquad {}^{*}T_{\beta} \;\to\; {}^{(+)}T_{\beta\mu} + \varpi^{(-)}T_{\beta\mu}\,,$$

$${}^{*}({}^{*}G_{\beta}) \;\to\; {}^{(-)}G_{\beta\mu}\,, \quad {}^{*}({}^{*}T_{\beta}) \;\to\; -\left({}^{(-)}T_{\beta\mu} + \omega^{(+)}T_{\beta\mu}\right).$$

Each solution of (2.3) and (2.4) is a *Friedmann-Lemaitre-Roberston-Walker metric*

$$ {}^{(\pm)}\mathrm{d}s^2 = \mathrm{d}t^2 - {}^{(\pm)}a(t)^2 \, \frac{\mathrm{d}u^2 + u^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right)}{\left(1 + \dfrac{k\,u^2}{4}\right)^2}, \qquad (2.5) $$

where k is referred to as the *curvature index*: $\{-1, 0, 1\}$.

Ultimately, inspection shows that:

$$ \varpi = \frac{{}^{(-)}a^3}{{}^{(+)}a^3} \quad \text{and} \quad \omega = \frac{{}^{(+)}a^3}{{}^{(-)}a^3}, \quad \omega = \varpi^{-1}. \qquad (2.6) $$

3 Conclusions and outlook

According to the Cosmological Janus Model, mass and charge inversions simultaneously result from time reversal which grant the theory a particularly simple and exhaustive symmetry.

As a final point, let us emphasize that the *JCM* bi-metric scheme is far from being an arbitrary postulate as it proves consistent with the newest developments in astrophysics.

It is also formally sustained by a specific splitting of the *Riemann tensor* in two 2nd rank tensor field equations as shown in [10]. This 4th rank tensor theory eventually leads to the space-time of constant curvature (i.e. in vacuum). It thereby copes with the recent view suggesting that the laws of physics are invariant under the symmetry group of *De Sitter* space (maximally symmetric space), rather than the *Poincaré* group of Special Relativity [11–14].

Submitted on March 24, 2019

References

1. Petit J.-P., D'Agostini G. Negative mass hypothesis in cosmology and the nature of the dark energy. *Astrophysics and Space Sciences*, 2014, v. 354, 611–615.
2. Sakharov A.D. Cosmological models of the Universe with reversal a time arrow. *Soviet Physics JETP*, 1980, v. 52, issue 3, 689–693 (translated from: *ZhETF*, 1980, v. 52, 349–351).
3. Petit J.-P., D'Agostini G. Constraints on Janus Cosmological model from recent observations of supernovae type Ia. *Astrophys. Space Sci.*, 2018, v. 363, 139.
4. Debergh N., Petit J.-P., D'Agostini G. On evidence for negative energies and masses in the Dirac equation through a unitary time-reversal operator. *J. Phys. Comm.*, 2018, issue 2, 115012.
5. Marquet P. Lichnerowicz's Theory of Spinors in General Relativity: the Zelmanov Approach. *The Abraham Zelmanov Journal*, 2012, v. 5, 117–133.
6. Kramer D., Stephani H., Hertl E., Mac Callum M. Exact Solutions of Einstein's Field Equations. Cambridge University Press, 1979.
7. Straumann N. General Relativity and Relativistic Astrophysics. Springer-Verlag, Berlin, 1984.
8. Dirac P.A.M. General Theory of Relativity. Princeton University Press, 2nd edition, Cambridge University Press, 1975, p. 61.
9. Marquet P. Vacuum background field in General Relativity. *Progress in Physics*, 2016, v. 12, issue 4, 314–316.
10. Marquet P. On a 4th rank tensor gravitational field theory. *Progress in Physics*, 2017, v. 13, issue 2, 106–110.
11. Aldrovani R., Beltran Almeida J.P., Pereira J.G. Some implications of the cosmological constant to fundamental physics. arXiv: gr-gc/0702065.
12. Lev F.M. De Sitter symmetry and quantum theory. arXiv: 1110.0240.
13. Aldrovani R., Beltran Almeida J.P., Pereira J.G. De Sitter Special Relativity. 2007.
14. Inönü E., Wigner E.P. *Proc. Natl. Acad. Scien.*, 1953, v. 39, 510.

Evidence of Residual Strong Interaction at Nuclear-Atomic Level via Isotopic Shift in LiH-LiD Crystals

V. G. Plekhanov[1] and J. Buitrago[2]
[1] Informatika ja Arvutustehnika Instituut, Tallinn, Estonia
E-mail: vgplekhanov@gmail.com
[2] Department of Astrophysics of the University of La Laguna, Faculty of Physics, 38205, La Laguna, Tenerife (Spain)
E-mail: jbuitrag@ull.es

Artificial activation of the strong interaction by adding one neutron to the nucleus causes the global reconstruction of the macroscopic characteristics of solids. The experimental evidence of macroscopic manifestation of the strong interaction in the optical spectra of solids which differ by one neutron from each other (using LiD crystals instead LiH ones) is presented for the first time. As far as the electromagnetic and weak interactions are the same in both kind of crystals, it only changes the strong interaction, therefore the renormalization of the energy of electromagnetic excitations (electrons, excitons, phonons) is carried out by the strong nuclear interaction. The necessity to take into account some new residual inter-relations between strong and electromagnetic interactions are underlined. An interpretation of the isotopic shift caused by the addition of one neutron is also discussed. From the experimental value of the isotopic shift we obtain a residual strong coupling constant equal to 2.4680.

1 Introduction

To the present we have a clear picture about the different kind of interactions and their main scenarios: electromagnetic ones for the realm of atomic physics and strong interactions for nuclear physics [1, 2]. However, in this articles we would like to report about some new experimental evidence, together with a tentative theoretical interpretation, pointing towards some relationship between both kind of interactions, which seems to lead to a new understanding in which nuclear forces can reach outside the nucleon boundaries and manifest themselves at the atomic level, at least in the magnetic manifestation. In what follows we shall try to explain how residual strong like interactions can affect, via electronic excitations (electrons, excitons, phonons) through isotopic effects, the binding energy of the dielectrics LiH and LiD crystals [3].

Nowadays in text books and elsewhere the separation of electromagnetic and strong interactions is tacitly assumed. Our results shine a new light on some residual interaction (ultimately based in the character of magnetic forces, of electromagnetic or color origin, which by their very nature, are difficult to conceal within the elusive nucleon physical boundary) between both kind of forces which is experimentally manifested trough isotopic shift. We hope that the results that we report in this paper will give a new insight about the manifestation of nuclear forces, by isotopic shift, beyond the nuclear domain.

2 Experimental results

In this part we shall describe the results of the optical spectroscopy of isotope-mixed solids (see, also [3]). The apparatus used in our experiments has been described in several previous publications [4, 5]. For clarity, we should mentioned here that immersion home-made helium cryostat and two identical double-prism monochromators were used. One monochromator was used for the excitation and the other, which was placed at right-angle to the first, for analyzing the luminescence and scattering of light. In our experiments we investigated two kinds of crystals (LiH and LiD) which only differ by the addition of one neutron. In view of the high hygroscopy of the investigated samples, the crystals were cleaved directly in liquid (superfluid) helium in the cryostat bath [4]. This makes possible to prepare samples with a clean surface. We found no changes in the free-exciton luminescence or resonance Raman scattering (RRS) [5] spectra when a sample with such a surface was studied for periods lasting 15 hours. The crystals were synthesized from 7Li metal and hydrogen 99.7 per cent purity and deuterium of 99.5 per cent purity (see, e.g. [3, 5] and references therein). We should remind very briefly about the electronic excitations in solids. According to modern concept, the excitons can be considered [6] as the excitation of the N-particles system: An electron from the valence band of insulators (see Fig. 1) is excited into the conduction band.

The attractive Coulomb potential between the missing electron in the valence band, which can be regarded as a positively charged hole, and the electron in the conduction band gives a hydrogen-like spectrum with an infinitive number of bound state and ionization continuum. In this article we call the bound states of electron-hole (e-h) pairs exciton states (exc), while we refer to ionized e-h pairs as free carriers. However, the expression free carriers does not imply that the effect of the strong Coulomb forces between electronic excitation could be neglected. Thus, an exciton state can be built by appropriate superposition of e-h pairs, which in a simple two-band model for cubic crystal symmetry is

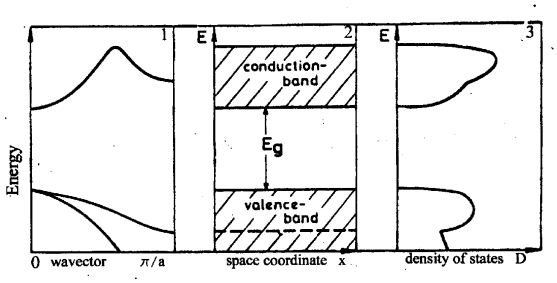

Fig. 1: Various possibilities to present the band-structure of homogeneous, undoped insulator (semiconductor). 1 - the dispersion relation, i.e. the energy E as a function of the wave vector, 2 - the energy regions of allowed and forbidden states as function of a space coordinate x and, 3 - the density of states (all curves are schematic ones).

given (for more details see [6]). As demonstrated some time ago [4] most low - energy electron excitation in LiH crystals are the large-radius excitons [6]. Exciton luminescence is observed when LiH (LiD) crystals are excited in the midst of the fundamental absorption. The spectrum of exciton photoluminescence of LiH crystals cleaved in liquid (superfluid) helium consists of a narrow (in the best crystals, its half-width is E $\leq$ 10 meV) phononless emission line and its broader phonon repetitions, which arise due to radiative annihilation of excitons with the production of one to five longitudinal optical (LO) phonons (see Fig. 2).

The phononless emission line coincides in an almost resonant way with the reflection line of the exciton ground state which is indication of the direct electron transition X_1 - X_4 of the first Brillouin zone [4].The lines of phonon replicas form an equidistant series biased toward lower energies from the resonance emission line of excitons. The energy difference between these lines in LiH crystals is about 140 meV, which is very close to the calculated energy of the LO phonon in the middle of the Brillouin zone and which was measured in (see, e.g. [3] and references therein). As we can see from Fig. 2 the photoluminescence spectrum of LiD crystals is largely similar to the spectrum of intrinsic luminescence of LiH crystals. The isotopic shift of the zero phonon emission line of LiH crystals equals 103 meV. There are, however, some related distinctions. Firstly the zero-phonon emission line of free excitons in LiD crystals shifts to the short-wavelength side on 103 meV. The second difference concludes in less value of the LO phonon energy, which is equal to 104 meV. Comparison of the experimental results on the luminescence and light scattering [3] in the crystals which differ by only one neutron is allowed to the main conclusion motivating this work: The addition of one neutron (using LiD crystals instead LiH ones) produce an unexpected increase of 103 meV in the exciton energy which seems rather difficult to explain within the conventional solid state physics scenario.

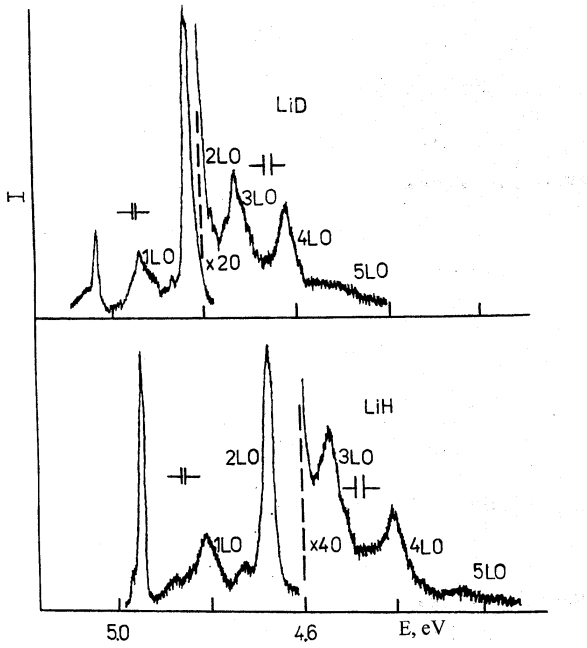

Fig. 2: Photoluminescence spectra of free excitons at 2 K in LiH and LiD crystals cleaved in superfluid helium.

3 Interpretation of the Isotopic Shift

We are used to find characteristic energies, mostly due to electrical interactions, of the order of one eV in the atomic and molecular scenarious. The reported experimental result of 0.103 eV emerging from magnetic-like interaction (magnetic forces are a factor v/c weaker than electric ones) is a surprising result pointing towards something that has not been observed before. The following comments, although tentative, pretend to give a plausible physical picture of new dynamical effects extending beyond the undefined borders of nucleons. From many experiments in QCD we know that direct forces between quarks are strong color analogues of electrostatic forces. However, in QCD, like in all gauge theories formulated within the context of special relativity, magnetic effects are unavoidable. In its more simple-minded description, electrostatic-like interactions between quarks have an origin and sink in the individual quarks confined in a nucleon. However, since quarks are not at rest, magnetic-like effects have to arise. The question we now ask is whether these effects should be limited to the inside-nucleon region or, perhaps, propagate outwards. In the absent of magnetic monopoles, magnetic force lines are closed. Moreover, magnetic fields are related to the SO(3) rotation group and its SU(2) covering group and it is not evident, at least in principle, if they couple only to ordinarily charged particles (remember that the SU(2) group is also contained in SU(3)). Consequently, in what follows, we shall consider, as an Ansatz, that magnetic color like

forces also couple to charged leptons. From the experimental results described before which arise by adding a single neutron to the LiH crystal:

$$LiH + n \longrightarrow LiD.$$

It seems that the 0.103 eV value should be regarded as an isotopic shift attributed to the magnetic moment of the charge-neutral neutron.

We are already familiar with the dipole-dipole magnetic interaction arising from the hyperfine splitting in the Hydrogen atom (for an adequate, to our purpose, study see [7]). The ground state wave function for the electron in the Hydrogen atom, including the spin part, is

$$\psi_0 = (\pi {a_0}^3)^{-1/2} e^{-r/a_0} |s>, \tag{1}$$

a_0 being the Bohr radius. We also need the energy of a magnetic dipole $\vec{m}_1$ in a magnetic field $\vec{B}$ produced by another dipole ($\vec{m}_2$) given by

$$H = -\vec{m}\,\vec{B}.$$

$$H = -\frac{1}{4\pi}\frac{1}{r^3}\left[3(\vec{m}_1\hat{r})(\vec{m}_2\hat{r}) - \vec{m}_1\,\vec{m}_2\right] - \frac{2}{3}(\vec{m}_1\,\vec{m}_2)\delta^3(\vec{r}). \tag{2}$$

As is well known, for s states with spherical symmetry the first term vanishes and only the second term involving a delta function contributes. This is essential as the wave function (1) has a finite value for $r = 0$ so that the energy comes out from a contact-interaction (see [7]). The magnetic dipole-dipole interaction can thus be treated as a perturbation. In first order perturbation theory:

$$E' = \int {\psi_0}^* H \psi_0 dV. \tag{3}$$

As mentioned, only the second term contributes giving:

$$E' = -\frac{2}{3} < \vec{m}_1\,\vec{m}_2 > |\psi_0(0)|^2 = -\frac{2}{3}\frac{1}{\pi {a_0}^3} < \vec{m}_1\,\vec{m}_2 > . \tag{4}$$

For the electron-proton we have two configurations according to the spin of both particles:

$$\vec{m}_1 = \gamma_p \vec{S}_p, \qquad \vec{m}_2 = -\gamma_e \vec{S}_e.$$

(γ: gyromagnetic ratio; $\gamma = (e/2m)g$, the g-factor being 2.0023 for the electron and 5.5857 for the proton.)

According to equation (4), we obtain for the triple and singlet states in Hydrogen, the energies

$$E'_t = \frac{1}{3}\frac{e^2}{a_0^3 m_e M_p} g_p = 1.4685 \times 10^{-6}\ \text{eV}$$

and

$$E'_s = -\frac{e^2}{a_0^3 m_e M_p} g_p = -4.4054 \times 10^{-6}\ \text{eV},$$

with a gap $\Delta E' = 5.874 \times 10^{-6}$, coincident with the hydrogen hyperfine splitting experimental result.

Similar calculations can be easily carried out for Deuterium (spin 1 and gyromagnetic ratio $g_d = 1.71$) with the results:

$$E'_{3/2} = 4.4980 \times 10^{-7}\ \text{eV},$$

$$E'_{1/2} = -8.9960 \times 10^{-7}\ \text{eV},$$

$$\Delta E'_d = E'_{3/2} - E'_{1/2} = 1.3494 \times 10^{-6}\ \text{eV}.$$

Turning now to the *Isotopic shift* issue, from the above values, we have four alternatives depending on the relative spins, however, as the lowest energy for both LiH and LiD is the corresponding to singlet states, we shall choose:

$$\Delta E = (E'_s)_H - (E'_{1/2})_D = -3.5058 \times 10^{-6}\text{eV}, \tag{5}$$

far from the experimental 0.103 eV. Next we shall assume that the experimental isotopic shift of 0.103 eV is the result of the onset of a residual strong interaction when the neutron is added, accordingly we do not modify $(E'_s)_H$ but modify $(E'_{1/2})_D$ in the following way: In Hydrogen the absolute value of the charge is the same so that in electric or magnetic interactions the coupling constant is $\alpha = e^2$. However, as the neutron do not have electric charge, in the dipole magnetic interaction the effective coupling constant can be defined through the transformation

$$\alpha = e^2 \longrightarrow (\alpha_s)_{\text{eff}} = e\, e_s. \tag{6}$$

The Bohr radius is thus modified:

$$a'_s = \frac{1}{e\, e_s}\frac{1}{m_e}.$$

From (4), it is easy to obtain

$$(E'_{1/2})_D = -\frac{4}{3} g_d \frac{(\alpha'_s)^4 m_e^2}{M_d}. \tag{7}$$

Inserting in (5) the 0.103 experimental value for ΔE and solving for α'_s, we obtain:

$$\alpha'_s = 0.1342,$$

and a *strong charge*

$$e_s = \frac{0.1342}{0.08542} = 1.5710,$$

leading to a strong coupling constant $e_s^2 = \alpha_s = 2.4680$. Quite large in comparison with the normal fine structure constant.

4 Conclusions

The experimental evidence of the macroscopic manifestation of strong nuclear interaction in optical spectra of solids which are differing by one neutron from each other has been presented for the first time. This evidence is based on two independent experimental results, which is directly seen from luminescence and reflection spectra. Our interpretation is based in the neutral charge of the neutron which in turn is responsible for the observed isotopic shift. We should be aware of the delicate interplay between solid state physics translated for a theoretical interpretation to the nuclear and subnuclear background which we have tried to accomplish in a way that could be regarded as somewhat tentative but unavoidable given the uncertainties laying in the strong magnetic-like interaction between nucleons and electrons.

Submitted on April 22, 2019

References

1. Henley E.M., Garcia A. Subatomic Physics. World Scientific Publishing Co., Singapore, 2007.
2. Griffiths G. Introduction to Elementary Particles. Wiley - VCH, Weinheim, 2008.
3. Plekhanov V.G. Isotopes in Condensed Matter. Springer, Heidelberg, 2013.
4. Plekhanov V.G. Experimental manifestation of the effect of disorder on exciton binding energy in mixed crystals. *Phys. Rev. B*, 1996, v. 53, 9558–9593.
5. Plekhanov V.G. Macroscopic manifestation of the strong nuclear interaction in the optical spectra of solids. In: *Proceedings of XXV International Seminar on Interactions of Neutrons with Nuclei*, ISINN - 25, Dubna, Russia, 2018, pp. 49–56.
6. Knox R.S. Theory of Excitons. Academic Press, New York - London, 1963.
7. Griffiths D. Hyperfine splitting of the ground state of hydrogen. *American Journal of Physics*, 1982, v. 50, no. 8, 698–703.

Can We Hide Gravitational Sources behind Rindler Horizons?

Michael Edward McCulloch

SoBMS, Plymouth University, Plymouth, UK. E-mail: mike.mcculloch@plymouth.ac.uk

When an object accelerates in one direction, a Rindler horizon forms in the opposite direction and information from behind it cannot reach the object. Here it is shown that it is possible to test for this effect since it predicts that if an object, say a disc, is rotationally accelerated by over $\sim 10^{10}\,\mathrm{m/s^2}$ then the Rindler horizon it sees should come close enough to hide part of the Earth and therefore it should not feel all the Earth's gravity. This effect could be detected by measuring the disc's weight.

1 Introduction

Hawking [1] showed that the strong gravity at the edge of a black hole produces an event horizon that can separate paired virtual particles leading to Hawking radiation and black hole evaporation. Fulling [2], Davies [3] and Unruh [4] showed that a similar effect occurs for accelerating objects in that a Rindler horizon [5] forms at a distance of c^2/a from the side they are accelerating away from (where c is the speed of light and a is the acceleration of the object). This horizon similarly produces radiation so that an accelerated object will perceive a warm background full of blackbody radiation whereas an unaccelerated body will see a cold background with no radiation. This is called Unruh radiation [4] and for typical accelerations it has too long a wavelength to be detectable, but it may have been observed coming from plasmons propagating at high acceleration around the surface of a gold nanotip [6].

McCulloch [7, 8] proposed a new model for inertia (called quantised inertia, or QI) that assumes that the inertia of an object is due to the Unruh radiation it sees when it accelerates. The Rindler horizon that appears in the opposite direction to its acceleration damps the Unruh radiation on that side of the object producing a radiation pressure differential that looks like inertial mass [8]. Also, when accelerations are extremely low the Unruh waves become very long and are also damped, this time in all directions, by the Hubble horizon (Hubble-scale Casimir effect). This leads to a new loss of inertia as accelerations become tiny. QI modifies the standard inertial mass (m) to a modified one (m_i) as follows:

$$m_i = m\left(1 - \frac{2c^2}{|a|\,\Theta}\right), \tag{1}$$

where c is the speed of light, Θ is twice the Hubble distance, $|a|$ is the magnitude of the relative acceleration of the object relative to surrounding matter. Eq. 1 predicts that for terrestrial accelerations (eg: $9.8\,\mathrm{m/s^2}$) the second term in the bracket is tiny and standard inertia is recovered, but in low acceleration environments, for example at the edges of galaxies (when a is tiny), the second term in the bracket becomes larger and the inertial mass decreases in a new way so that QI can predict galaxy rotation without the need for dark matter [9].

Putting Eq. 1 into Newton's second and gravity laws gives

$$F = ma = m\left(1 - \frac{2c^2}{|a|\,\Theta}\right) = \frac{GMm}{r^2} \tag{2}$$

and finally

$$a = \frac{GM}{r^2} + \frac{2c^2}{\Theta}\,. \tag{3}$$

This predicts cosmic acceleration (the new second term) without the need for dark energy [7]. In this paper this same result is derived a different way, simply using Ernst Mach's attitude that "what cannot be observed does not exist". It is argued that, since Rindler horizons are boundaries for information, then sources of gravity behind them disappear from the point of view of the accelerated object. It is shown here that this effect predicts cosmic acceleration, given the known baryonic mass of the cosmos, and may allow us to hide gravitational sources behind horizons producing new kinds of thrust.

2 Method

If we consider a photon travelling at the speed of light in the centre of its own Hubble sphere (see Fig. 1). Due to the impossibility of any light from the left hand side of the cosmos catching up to the photon, we can say that, as far as the photon knows, there is no mass there at all. All the mass is hidden by the Rindler horizon. Therefore, there is a gravitational imbalance as the photon can be aware of a lot of matter in front of it in the direction of its acceleration, but nothing behind. We can calculate this gravitational acceleration as follows

$$a = \frac{GM}{r^2}\,. \tag{4}$$

We can assume from standard geometry that the centre of mass of the semi-sphere in front of the photon is 3/8ths of the radius away, and the radius and baryonic mass of the cosmos are estimated to be 4.4×10^{-10} m and $10^{52\pm1}$ kg, so

$$a = \frac{6.67 \times 10^{-11} \times 10^{52\pm1}}{(3/8 \times 4.4 \times 10^{26})^2} = 2.45 \times 10^{-11\pm1}\mathrm{m/s^2}. \tag{5}$$

The predicted acceleration (given the error bars) agrees with the observed cosmic acceleration and with the critical acceleration below which galactic dynamics deviate from Newton: $2 \times 10^{-10}\mathrm{m/s^2}$.

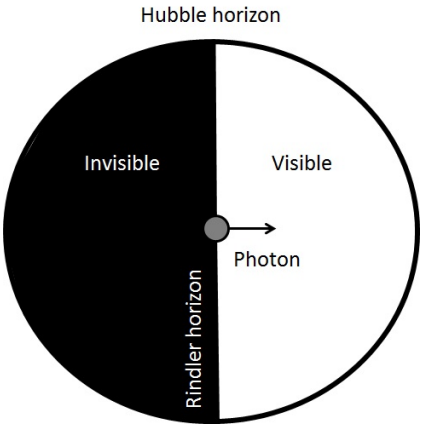

Fig. 1: A schematic showing the Hubble horizon (as a black circle). A photon (the central grey circle) moves rightwards at the speed of light, so it has a Rindler horizon passing through it, and no information from the black-shaded volume can get to it. This means that, following Mach, the gravitational mass from that black region is irrelevent and the gravitational pull from the right hand half of the cosmos now dominates, causing an acceleration which predicts the cosmic acceleration.

3 A test

If we consider a spinning disc, then every particle within it is accelerating towards the spin axis, and each particle perceives a Rindler horizon that is outside the disc. As the rotational acceleration is increased the horizon moves closer to the spin axis. What would happen if the horizon was closer than the Sun or the Earth? Would this hide their gravitational effect from the point of view of the accelerated object? (see an earlier brief discussion of this in [10]).

To calculate the spin rate required to pull the Rindler horizon in closer than a distance d_R we assume a disc of any material of radius r, spinning at R rpm (rotations per minute). The centripetal acceleration (a) at different radii (r) of the disc is given by

$$a = \frac{v^2}{r} = \frac{(2\pi rR/60)^2}{r} = \frac{4\pi^2 rR^2}{3600}\,, \tag{6}$$

where the 60 comes from the number of seconds in a minute. The Rindler horizon forms in the direction opposite to the acceleration at a distance given by

$$d_R = \frac{c^2}{a}\,. \tag{7}$$

We can now substitute Eq. 6 in Eq. 7

$$d_R = \frac{3600\,c^2}{4\pi^2 rR^2} = \frac{900\,c^2}{\pi^2 rR^2}\,. \tag{8}$$

Eq. 8 shows the distance of the Rindler horizon (d_R) for a particle within a disc spinning at R rpm and at a radius r from the spin centre. It shows that the faster the disc spins (R increases) the distance to the Rindler horizon decreases very rapidly and the Rindler horizon is closer for particles at the disc's edge (when r is large).

4 Results & discussion

Eq. 8 can be rearranged to calculate the rotation rate R (in rpm) needed to bring the Rindler horizon closer than a body a distance d_R away

$$R = \sqrt{\frac{900\,c^2}{\pi^2 r d_R}}\,. \tag{9}$$

The following table shows the object to be hidden by the Rindler horizon in the first column. The second column shows its distance (d) away from a lab on the Earth's surface. The third column shows the acceleration needed, in a linear sense, to hide the object. The fourth column shows the rpm required for a spinning disc to achieve that acceleration, at a radius of 0.1 m. The fifth column shows the gravitational acceleration ($a_g = GM/d^2$) produced by that object that will disappear and affect the dynamics of the disc (but only those parts of it above the critical acceleration).

Object	Distance	a	rpm	a_g
		Eq. 7	Eq. 9	
		(m/s^2)		(m/s^2)
Sun	1 AU	600,000	23 k	0.006
Earth	6371 km	1.43×10^{10}	3589 k	9.8

Table 1: The Table shows for two objects (column 1), the distances from a lab on the Earth's surface to the object (column 2), the accelerations needed to hide the object behind Rindler horizons (column 3), the rpm needed for that acceleration for a disc at a radius of 0.1 m (column 4) and the acceleration exerted by the object on the disc (column 5).

The rotation required to hide the Sun should be achievable since gyroscopes often have rotation rates of 30,000 rpm and medical centrifuges can spin at 100,000 rpm. The rotation rate required would be lower for a larger disc. Of course, only the part of the disc that has an acceleration vector pointing away from the Sun (the Sunward side) would feel the disappearence of the Sun's effect, including its gravitational force. The gravitational acceleration due to the Sun is $GM_\odot/r^2 = 0.006\,m/s^2$ (this is 0.06% of g). The Sun's width in the sky is about half a degree so only an area of about 1/(360*2) of the disc would be affected and then also only the area of the disc outside the radius of 0.1m. So if the disc was 0.2m in radius the affected area would be the total area times (1/720)×(3/4). Therefore, the average acceleration for the whole disc would be $0.006 \times (1/720) \times (3/4) = 6.25 \times 10^{-6} m/s^2$.

From a practical point of view it would be far more useful to hide the Earth's gravity since then launching objects would become easier. The acceleration required to do so: 1.43×10^{10} (see Table 1), has just been achieved for the first time by [11] who spun a microscopic sphere of radius $r = 4\times10^{-6}$ m using circularily polarised light to suspend and rotate it in vacuo at $R = 6 \times 10^8$ rpm. This is an acceleration, using Eq. 6 of $1.58 \times 10^{10}\mathrm{m/s^2}$ which agrees with the acceleration needed to pull the Rindler horizon close enough to hide the Earth's gravity (Table 1, column 3).

5 Conclusion

It is proposed here that Rindler horizons have physical consequences beyond their effects on light: they are able to hide gravitational sources.

It is shown that assuming that gravitational sources can be hidden in this way, predicts the cosmic acceleration.

The effect could be tested using discs with extreme spins, which should break free from distant gravitational sources.

Acknowledgements

Many thanks to J. Lucio for proof-reading this manuscript, and to DARPA for funding grant HR001118C0125.

Submitted on May 5, 2019

References

1. Hawking S. Black hole explosions. *Nature*, 1974, v.248, 30.
2. Fulling S.A. Nonuniqueness of Canonical Field Quantization in Riemannian Space-Time. *Phys. Rev. D.*, 1973, v.7, 2850.
3. Davies P.C.W. Scalar production in Schwarzschild and Rindler metrics *J. Phys. A.*, 1975, v.8, 609.
4. Unruh W.G. Notes on black hole evaporation. *Phys. Rev. D.*, 1976, v.14, 870.
5. Rindler W. Relativity, special, general and cosmological. Oxford University Press, 2001.
6. Smolyaninov I.I. Photoluminescence from a gold nanotip in an accelerated reference frame *Physics Letters A*, 2008, v.372, 7043–7045.
7. McCulloch M.E. Modelling the Pioneer anomaly as modified inertia. *MNRAS*, 2007, v.376, 338.
8. McCulloch M.E. Inertia from an asymmetric Casimir effect. *EPL*, 2013, v.101, 59001.
9. McCulloch M.E. Galaxy rotations from quantised inertia and visible matter only. *Astrophys. & Space Sci.*, 2017, v.362, 149.
10. McCulloch M.E. Physics from the edge: a new cosmological model for inertia. World Scientific, 2014.
11. Arita Y., Mazilu M. and Dholakia K. Laser-induced rotation and cooling of a trapped microgyroscope in vacuum. *Nature Communications*, 2013.

A Mathematical Definition of "Simplify"

Craig Alan Feinstein

2712 Willow Glen Drive, Baltimore, Maryland 21209. E-mail: cafeinst@msn.com

Even though every mathematician knows intuitively what it means to "simplify" a mathematical expression, there is still no universally accepted rigorous mathematical definition of "simplify". In this paper, we shall give a simple and plausible definition of "simplify" in terms of the computational complexity of integer functions. We shall also use this definition to show that there is no deterministic and exact algorithm which can compute the permanent of an $n \times n$ matrix in $o(2^n)$ time.

1 Introduction

In 2013, the author asked the following quesiton titled "Is there a 'mathematical' definition of 'simplify'?" on the popular mathematics website MathOverflow.net [1]:

> "Every mathematician knows what 'simplify' means, at least intuitively. Otherwise, he or she wouldn't have made it through high school algebra, where one learns to 'simplify' expressions like $x(y + x) + x^2(y + 1 + x) + 3(x + 3)$. But is there an accepted rigorous 'mathematical' definition of 'simplify' not just for algebraic expressions but for general expressions, which could involve anything, like transcendental functions or recursive functions? If not, then why? I would think that computer algebra uses this idea."

The answers there indicated that even though every mathematician knows intuitively what "simplify" means, there is still no universally accepted definition of "simplify". In fact, one of the answers (by Henry Cohn) indicated that "In full generality, there provably isn't any method for complete simplification". (He was referring to elementary functions of a real variable.) In this paper, we shall give a simple and plausible definition of "simplify" in terms of the computational complexity of integer functions. We shall also use this definition to show that there is no deterministic and exact algorithm which can compute the permanent of an $n \times n$ matrix in $o(2^n)$ time.

2 A definition of "simplify"

Consider the following definition of "simplify":

Definition: An algebraic expression (recursive or non-recursive) for a function $f : \mathbb{Z} \to \mathbb{Z}$ *cannot be simplified* if there is no other algebraic expression for f which can be computed faster.

For example, the expression $xw+yz+xz+yw$ can be simplified to $(x + y)(w + z)$, since computing $(x + y)(w + z)$ takes only one multiplication and two additions, while computing $xw + yz + xz + yw$ takes four multiplications and three additions. And we can also see clearly that the expression $(x + y)(w + z)$ cannot be simplified.

As another example, let $f : \mathbb{Z} \to \mathbb{Z}$ be the function which satisfies the recursive formula, $f(n) = f(n-1) + 1$ and $f(0) = 0$. This recursive formula can be simplified to $f(n) = n$, since computing the recursive formula for f takes $\Theta(n)$ time, while computing the formula $f(n) = n$ is trivial. And the formula $f(n) = n$ clearly cannot be simplified.

And let $f : \mathbb{N} \to \mathbb{N}$ be the function which satisfies the recursive formula, $f(n) = f(n-1) + f(n-2)$ and $f(1) = f(2) = 1$, the Fibonacci sequence. This recursive formula can be simplified, since it is possible to prove that $f(n)$ equals $\phi^n/\sqrt{5}$ rounded to the nearest integer, where $\phi = (1 + \sqrt{5})/2$, which can be computed exponentially faster than the recursive formula can be computed [4].

3 Computing the permanent of a matrix

Let $A = (a_{ij})$ be a matrix of integers. The permanent of A is defined as:

$$\mathrm{perm}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)},$$

where S_n is the symmetric group [5]. The fastest known deterministic and exact algorithm which computes the permanent of a matrix was first published in 1963 and has a running-time of $\Theta^*(2^n)$ [3]. It is still considered an open problem by the mathematics and computer science community whether this time can be beaten. Now consider the following theorem and proof, which we shall discuss afterwards:

Theorem: There is no deterministic and exact algorithm which can compute the permanent of an $n \times n$ matrix in $o(2^n)$ time.

Proof: For any row i, the permanent of matrix A satisfies the recursive formula

$$\mathrm{perm}(A) = \sum_{j=1}^{n} a_{ij} \cdot \mathrm{perm}(A_{ij}^{\#})$$

and $\text{perm}([a_{11}]) = a_{11}$, where $A^{\#}_{ij}$ is the $(n-1)\times(n-1)$ matrix that results from removing the i-th row and the j-th column from A. This formula cannot be simplified, so the fastest algorithm for computing the permanent of a matrix is to apply this recursive formula to matrix A. Since this involves recursively evaluating the permanent of $\Theta(2^n)$ submatrices of A, each corresponding to a subset of the n columns of A, we obtain a lower bound of $\Theta(2^n)$ for the worst-case running-time of any deterministic and exact algorithm that computes the permanent of a matrix. □

At first, this proof makes sense intuitively, but if one thinks about it a little more, one might become skeptical, since one could argue the same for the determinant of a matrix, that there is no deterministic and exact algorithm which can compute the determinant of an $n \times n$ matrix in $o(2^n)$ time (which is known to be false) - for any row i, the determinant satisfies the recursive formula

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot \det(A^{\#}_{ij})$$

and $\det([a_{11}]) = a_{11}$, which is almost the same as the recursive formula for the permanent of a matrix.

However, there is a big difference between the two recursive formulas: There are negative signs in the formula for the determinant, so it is not inconceivable that one might be able to cancel most of its terms out, if one is clever. And in fact this is the reason why it is possible to compute the determinant of a matrix in polynomial-time: If one performs elementary row operations on matrix A with pivot $a_{11} \neq 0$, converting it to a matrix B with zeroes in the last $n-1$ entries of column 1, then the determinant of A will equal the determinant of B and we will also obtain a simpler formula for the determinant:

$$\det(A) = a_{11} \cdot \det(B^{\#}_{11}).$$

This trick ultimately leads to a polynomial-time algorithm for computing the determinant of a matrix, if one applies it recursively to the matrix $B^{\#}_{11}$, exchanging rows when necessary.

However, in the case of the permanent of a matrix, no trick like this is possible, since there are only positive signs in its formula. To gain some insight as to why this is so, consider the following analogy: Suppose we want to subtract two large positive numbers with a tiny difference, say $a = 12,345,678,907$ and $b = 12,345,678,903$. One could compute a minus b by applying the normal subtraction procedure that one learns in elementary school to each digit of these two numbers, but one does not have to do this; if we let $c = 12,345,678,900$, then we will obtain the same answer by computing $(a - c)$ minus $(b - c)$, which amounts to subtracting only the last digits of each number, 7 minus 3. But there are no short-cuts like this for adding a and b, since none of their digits can be cancelled out. And for this same reason, it is possible to cancel out lots of terms in the formula for the determinant but not in the formula for the permanent, as the elementary row operations which are performed on matrix A when computing its determinant via the algorithm described above are analogous to subtracting c from both a and b.

But then one might ask, "The proof above said 'This formula cannot be simplified'. But how can I be sure of this?" The answer to this question is that we know that the above recursive formula for the permanent cannot be simplified, because we have tried every possible way to simplify it and saw that each way fails: To be specific, we tried to multiply the factors, a_{ij} and $\text{perm}(A^{\#}_{ij})$, of the summands together, but we failed since the two factors are completely independent from one another. And we tried adding the summands together, but we also failed since the factors a_{ij} found in each summand are completely independent from one another and are also completely independent from each $\text{perm}(A^{\#}_{ij})$; furthermore, we found that since $\text{perm}(A^{\#}_{ij})$ is different in each term, it is impossible to use the distributive law to decrease the computational complexity of the recursive expression. And finally, we noticed that the row choice of i is irrelevant in the recursive formula for the permanent, so no choice of i is better than any other choice. What other things are there to try that could possibly make the expression simpler? Nothing, since we have already considered every mathematical operation in the recursive formula for the permanent. Therefore, the recursive formula for the permanent cannot be simplified, i.e., it has the best computational complexity of any algebraic expression for the permanent of a matrix.

This type of reasoning is not new or foreign; it is essentially the same type of reasoning that a high school math student uses to simplify algebraic expressions. Also note that only if one is careful in one's analysis and considers every possible way to simplify an algebraic expression can one prove that an algebraic expression indeed cannot be simplified; merely claiming that an algebraic expression cannot be simplified does not make it so. But sometimes it is so obvious that an algebraic expression cannot be simplified that writing down a full explanation of this is unnecessary. Also, it turns out that one can use similar reasoning to prove that there is no deterministic and exact algorithm which solves the Traveling Salesman Problem in polynomial-time [2].

4 Conclusion

While everyone in the mathematics community understands intuitively what "simplify" in mathematics means, there is still no universal definition of "simplify". In this paper, we have defined "simplify" in terms of the computational complexity of an integer function and have shown that this definition can be used to prove that there is no deterministic and exact algorithm which can compute the permanent of an $n \times n$ matrix in $o(2^n)$ time.

Submitted on May 11, 2019

References

1. Feinstein C.A. Is there a 'mathematical' definition of 'simplify'? https://mathoverflow.net/q/126519/7089
2. Feinstein C.A. The Computational Complexity of the Traveling Salesman Problem. *Global Journal of Computer Science and Technology*, 2011, v. 11, Issue 23, 1–2. https://arxiv.org/abs/cs/0611082
3. Ryser H.J. Combinatorial Mathematics. Carus Math. Monograph No. 14, 1963.
4. Fibonacci Number. *MathWorld — A Wolfram Web Resource.* http://mathworld.wolfram.com/FibonacciNumber.html
5. Permanent. *MathWorld — A Wolfram Web Resource.* http://mathworld.wolfram.com/Permanent.html

Science's Dilemma – a Review on Science with Applications

Linfan Mao

Chinese Academy of Mathematics and System Science, Beijing 100190, P.R. China
Academy of Mathematical Combinatorics & Applications(AMCA), Colorado, USA
E-mail: maolinfan@163.com

Actually, different views result in different models on things in the universe. We usually view a microcosmic object to be a geometrical point and get into the macrocosmic for finding the truth locally which results in a topological skeleton or a complex network. Thus, all the known is local by ourselves but we always apply a local knowledge on the global. *Whether a local knowledge can applies to things without boundary*? The answer is negative because we can not get the global conclusion only by a local knowledge in logic. Such a fact also implies that our knowledge on a thing maybe only true locally. *Could we hold on the reality of all things in the universe globally*? The answer is uncertain for the limitation or local understanding of humans on things in the universe, which naturally causes the science's dilemma: it gives the knowledge on things in the universe but locally or partially. Then, *how can we globally hold on the reality of things in the universe? And what is the right way for applying scientific conclusions, i.e., technology*? Clearly, different answers on these questions lead to different sciences with applications, maybe improper to the universe. However, if we all conform to a criterion, i.e., the coexistence of human beings with that of the nature, we will consciously review science with that of applications and get a right orientation on science's development.

1 Introduction

As is known to all that being is nature. Science discovers rulers on things existed in the universe with observable physical evidence. It is a systematic knowledge on the universe in the view of human beings. However, it enables human beings coexistence with the universe thousand million years. Today, it is the time to review science's function on reality of things in the universe with speculation on questions for science. For example, *does the science hold with the universe globally, or only partially*? And *what is the right application of science*? All the answers will push forward science, and establish a right view on its applications.

2 Nature's laws

Science is established on an assumption that "*the universe is operating in order*" which implies the existence of natural laws, i.e., the inherent law on the existence and motion of things in the universe but independent on humans. This assumption is general accepted by scientific community or human beings without questions. Now, a more basic but philosophical question in front of humans is that *could we really holds on natural laws without artificial conditions*? And furthermore, *is human's ability with or without boundaries*? Although there exist certain differences in the eastern and western cultures but the answer is the same, i.e., we can only stand in awe of and never destroy the nature, such as the *Platonism* in Plato's Dialogues: "*the universals exist independently of particulars*", and the *Tao* and *Name* in Tao Te Ching: "*the Tao experienced is not the eternal Tao, the Name named is not the eternal Name; the unnamable is the eternally real and naming is the origin of all particular things*". All of these views conclude that the known natural laws are understood by human beings ourself. They are only laws in our eyes, maybe not the really natural laws.

How do we understand the reality or establish the knowledge on a thing T in the universe? We assume there is an abstract T defined by a conception, i.e., *name* distinguished from other things and usually identified T with known characters, gradually little by little and from time to time. For example, let $\mu_1, \mu_2, \cdots, \mu_n$ be the known on T and $\nu_i, i \geq 1$ unknown characters at a time t. Then, T is understood by [1]

$$T = \left(\bigcup_{i=1}^{n}\{\mu_i\}\right)\bigcup\left(\bigcup_{k\geq 1}\{\nu_k\}\right), \tag{1}$$

a *Smarandache multispace* [2] or *parallel universe* [3] in logic at the time t on its connotation and extension, which also reveals the diversity or complexity on the reality of things T. Then, *what is thing T and what is its reality*? Philosophically, the reality of a thing T is nothing else but the state characters (1) of existed, existing or will existing things whether they can be or not observable or comprehensible by human beings at time t. Thus, we can only hold on T by its an asymptotic $T^\circ = \bigcup_{i=1}^{n}\{\mu_i\}$ at time t, and deeply convince that $T^\circ \to T$ if the time $t \to \infty$. This is the essential notion that natural laws can be understood, i.e., establishing science of humans.

3 Science's limitation

As humans enter the 21st century, science has made great achievements both in theory and its applications. It greatly improved the ability to respond of natural disasters, brings more and more conveniences to human life. In fact, science

is the systematic knowledge with continuously improvement on asymptotically natural laws dependent on observation and speculation of ourself, maybe with the aid of instruments. Certainly, different standings for the observer will result in different observations, i.e., the macrocosmic or the microcosmic which result in different speculating models.

3.1 Macrocosmic object

A macrocosmic object is large enough to be visible by the naked eyes of humans. For knowing the behavior of macrocosmic objects, the observer only needs to stand out of the observing object, holds on the overall situation, i.e., its outside behavior, particularly, planetary motion which establishes classical mechanics. It should be noted that the thinking pattern of classical mechanics is essentially

$$\begin{array}{lll} \text{Macrocosmic Object} & \overset{Abstract}{\longrightarrow} & \text{Particle} \\ & \overset{Abstract}{\longrightarrow} & \text{Geometrical Point} \end{array}$$

with 2 assumptions, i.e., ① there exists an abstract geometrical space $\mathbb{R}^3\times\mathbb{R}$ in the universe, and ②, all physical quantities can be accurately measured by humans.

As is known to all, the classical mechanics is applying only to those of objects A moving at low speeds, characterizing an object of quality m by a pair $\{\mathbf{x}, \mathbf{v}\}$, where $\mathbf{x}$ is the coordinates of A with a directed velocity $\mathbf{v}$ at points $\mathbf{x}$. For example, if A moves in a conservative field with potential energy $U(\mathbf{x})$, then the force acting on A is $\mathbf{F} = -\frac{\partial U}{\partial \mathbf{x}} = m\ddot{\mathbf{x}}$ by the second law of Newton, and generally, the Euler-Lagrange equations [4]

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} = 0, \quad 1 \le i \le n \tag{2}$$

in $\mathbb{R}^n$ for the Lagrangian $L = T - U$ of A, where $T(\mathbf{x})$ is the moving energy of A.

Although it is on macrocosmic objects, the classical mechanics found the intrinsic essence of motion, i.e., force. For example, Newton realized the gravity by an apple fell on his head from a tree and proposed the law of universal gravity $F = G\frac{M_1M_2}{R^2}$ between 2 bodies with masses M_1 and M_2 respectively, where R is the distance of the 2 bodies and G the constant of universal gravity. Although Newton's law is an approximation of gravity, it is useful in aerospace engineering. By this law, we have known the cosmic speeds surround the earth, escaped from the earth or the solar system are respectively 7.9 km/s, 11.2 km/s and 16.7 km/s which enables launching satellites for space exploration and communication of humans.

By the general relativity, i.e. *all the laws of physics take the same form in any coordinate system* and the equivalence principle, i.e., *there are no difference for physical effects of the inertial force and the gravitation in a field small enough*, Einstein presented the gravitational equations

$$R_{\mu\nu} - \frac{1}{2}\,g_{\mu\nu}R = \kappa T_{\mu\nu}, \tag{3}$$

where, $T_{\mu\nu}$ is the energy-momentum tensor, $R_{\mu\nu} = R^{\alpha}_{\mu\alpha\nu} = g^{\alpha\beta}R_{\alpha\mu\beta\nu}$, $R = g^{\mu\nu}R_{\mu\nu}$ are respectively the *Ricci tensor*, *Ricci scalar curvature* and $\kappa = \frac{8\pi G}{c^4} = 2.08\times10^{-48}\ \mathrm{cm^{-1}\ g^{-1}\ s^2}$.

Clearly, an immediate application of Einstein's gravitational equations is on the spacetime structure of the universe. For example, if it is in vacuum, i.e., $T_{\mu\nu} = 0$, the Einstein gravitational equations were solved due to the assumption of spherically symmetric distribution of matters and get the Schwarzschild metric $d^2s = g_{\mu\nu}dx^{\mu\nu}$ by

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{dr^2}{1-Kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\varphi^2\right)\right] \tag{4}$$

with $g_{tt} = 1$, $g_{rr} = -\frac{R^2(t)}{1-Kr^2}$, $g_{\phi\phi} = -r^2R^2(t)\sin^2\theta$, which also predicts the existence of black hole in the universe. Combining the cosmological hypothesis, i.e., *there are no difference at different points and different orientations at a point of the universe on the metric* 10^4 *l.y.*, Friedmann presented the *Standard Model on Universe* which resulted in the *Big Bang theory* in thirties of the 20th century and the scenario of the universe, i.e., it has a beginning.

Certainly, classical mechanics successfully explains a few astronomical phenomena, particularly, the planetary motion laws in front of humans thousands years. However, it is only an interpreting on the extrinsic behaviors but difficult on the internal cause, the basis for the change of objects. Today, we have known there is an additional assumption on a moving object in classical mechanics, i.e., all parts of the object are moving in coherence or synchronization. It is this assumption that can not enables humans globally understanding the nature of objects because the non-coherence, i.e., contradiction is the general but the coherence is the special, and all of us know that it is the contradiction or non-coherence pushes forward the change of things. Thus, holding on the nature of an object enables human's observation entering the microcosmic world with the aid of instruments and exploring microcosmic behavior of objects, i.e., microcosmic particles.

3.2 Microcosmic particle

A matter can be always divided into submatters, then subsubmatters and so on. A natural question on this subdividing is *whether or not it has a terminal point*? The answers are the same both for the Easterners and Westerners. For example, the ancient Chinese had a notion that *everything is composed by five elements, i.e., metal, wood, water, fire and soil* and also, the notion that *everything is composed entirely of various imperishable, indivisible elements*, i.e., there exist atoms in Atomism of Leucippus and Democritus, which finally results in the structure theory on matters. Today, it is already

a public knowledge that all matters are made up of atoms, i.e., microcosmic particles composed of nucleus with electrons. There are 118 atoms known by humans which consist of known matters on the earth. Generally, we understand a matter by the composite of elementary particles with a thinking pattern following

$$\text{Matter} \xrightarrow{Decompose} \text{Microcosmic Particles} \xrightarrow{Abstract} \text{Complex Network}$$

where the complex network is an inherit structure of the matter on microcosmic particles and different subjects discuss microcosmic behaviors of particles.

3.2.1 Physics

Clearly, the subdividing on a matter can be done infinite times just like the claim that "*it will be never exhausted if you cut half on a stick each day*" on World Chapter of Zhuang Zi in the ancient China. However, this process can not be applied to hold on matters because the life time of a human is not infinite. The motivation of particle physics is to determine the nature of irreducibly smallest detectable particles [5], called *elementary particles* such as those of *fundamental fermions* including quarks, antiquarks, leptons, antileptons and *fundamental bosons* including gauge bosons, Higgs boson and the *fundamental interactions* for explaining their behavior and then, the origin of the universe. Certainly, there are also unmatters between a matter and its antimatter which is partially consisted of matter but others antimatter [6]. However, the behavior of a microcosmic particle maybe indefinite. It is this character that results in humans characterizing microcosmic particles by wave function, a complex-valued probability amplitude.

In the non-relativistic quantum mechanics, we know that the wave function $\psi(t, x)$ of a particle of mass m obeys the Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + U \tag{5}$$

with the Planck constant $\hbar = 6.582 \times 10^{-22}$ MeVs and the potential energy U of field which characterizes the behavior of microcosmic particles.

Certainly, physics has promoted the progress of human society with the deeply understanding of matters from the macrocosmic to the microcosmic such as those of the applications of steam engine, the electricity with radio communication, nuclear energy, laser, electronic computer technology and so on. We seriously conclude that if there were no the development of physics, there would be no other sciences and no modern life of humans.

3.2.2 Chemistry

According to the notion that chemical compounds are not a random but rather definite one of atoms, the chemistry determines the composition, structures and properties of matters, particularly on atomic and molecular systems for the pattern and multiplicity of bonding between atoms in a molecule for explaining chemical reactions of matters. Although physics and chemistry are both on the structure of matters, the chemistry discusses the coarse-graining particles, i.e., atoms and molecules with chemical dynamics on rates of chemical reactions, but not on the fermions, bosons and their interactions.

Chemistry is beneficial for humans with a core topic, i.e., *how to create new matters to meet the needs of our daily life* in its developing. If there were no chemistry there would be no modern life of humans. For example, the chemical fertilizer increases the production of crops for maintaining the survival of population, the chemical pesticides kill insects harmful to crop growth, the medicines heals the sick with life extended, the plastics and synthetic fibers are used both in industrial and consumer products such as those of keyboard, mouse, plastic cup, slippers in our daily life, machinery, electronic appliances, automobile products, and furthermore, the dynamite, bombs and missile in military. None of them is not the application of chemistry.

3.2.3 Biology

Historically, biology is the oldest subject with the development of science in natural philosophy because humans ourself is also one specie of livings on the earth. Observation enables humans held on the elementary rotate regulation of plants on seasons, i.e., *spring germination with harvest or leaves fallen in autumn* and the reproduction regulation of humans and other animals such as "*pregnancy* 10 *months with childbirth in a day*" of humans, enables humans living together with the nature in about 5 million years. Certainly, the birth and the death are the two sides but all of us wish to hold on the laws of livings with production, the central issues of biology.

According to the notion that the basic unit of life is cell, the basic unit of heredity is genes and all life on the earth changes and develops through evolution, biology is such a science that on the life and living organisms respectively at molecular, cell, genes and heredity with variation levels and the process of grow and developing. Certainly, all major issues in the developing of humans society such as those of population growth, food safety, health, environmental pollution and resource depletion have a closely relationship with the life sciences. The project on human genome puts into effect with development will enables humans understanding the mechanism of growth, development, physiological activities and pathogenesis of diseases, which provides methods of prevention and control strategies on diseases of human bod-

ies, particularly, the gene and cell engineering. For examples, the transgentic technology can improves the crops resistance to insects for solving the pesticide residue problem and improving the quality of agricultural products; the antigen gene can applicable to the production of edible crop vaccine; the animal organs can be transplanted into a human body to play the role of such human organs, the cloning technology can detectable the fetal genetic defects, treats the injury of nervous system, achieves the asexual reproduction and saves the endangered species; the gene editing can correctable the defective gene for the treatment; the gene engineering can applicable to the environmental governance for recycling the pesticides and industrial wastes, and the large-scale animal cell culture can produces vaccines, breeds good varieties, detects the difference between virus strains and identifies the bacterial species for disease treatment, . . . , etc.

3.3 Science's limitation

However, all scientific conclusions of humans hold on conditions. *Is there such a scientific conclusion constraint without conditions*? The answer is negative both in theoretical and experimental sciences because of the boundary of humans. For example, all theorems are true with an obvious or implication that "*if p then q*" in mathematics. Even if the elementary conclusions $1 + 1 = 2, 1 \times 2 = 2$ known by pupils is such one only because they are implicit, i.e., "*if* $1, 2 \in (\mathbb{Z}; +, \times)$ *then* $1 + 1 = 2$ *and* $1 \times 2 = 2$", where $(\mathbb{Z}; +, \times)$ is the integer ring.

Similarly, we have known that sciences such as those of physics, chemistry, biology on a matter T by the macrocosmic are on its external behaviors with an additional assumption that all of its microcosmic particles are synchronous because it is abstracted to be a point with relatively external motion in space. We conclude that force is the internal factor of motion, creates new matters by chemistry and apply bionics to enrich human's living by simulating other creatures to conform the nature.

All of us known that the external causes operates through internal factor but a scientific conclusion on a matter T by the microcosmic is only partial or local nature because T is a complex system or a complex network in the thinking pattern. Until today, we lack of effective methods, even lack of such a mathematics on complex network or complex system which can not enables us hold on the whole matter T in theory unless all its microcosmic particle are in synchronization. So, we have only an incomplete or non-comprehensive science for things in the universe which is the limitation of human's science, an immediately conclusion of formula (1), i.e., the boundary of humans. In this case, we can hardly conclude that a scientific conclusion is true in the whole universe because it is understanding only by humans ourself, an intelligent creature happily born on the earth and it is a conclusion on known or unknown conditions.

4 Science's dilemma

4.1 Reality

Science's function is to understand the reality of things T, i.e., their state of existed, exists or will exist in the universe, whether or not they are observable or comprehensible by humans. However, this is difficult from the limitation of science because all scientific conclusions of humans are true constraint with conditions. They are locally or partially true, not freely with conditions or on the whole universe because we hold it little by little with an asymptotic T° of T, not T itself at a time t by formula (1). Usually, the physical laws are characterized by differential equations. Even for physical reality with differential equations, there are also 3 simple but basic questions should be answered.

Question I *Could a special solutions be applied to the whole universe*?

The answer of Question I is obviously negative unless the equations have a unique solution but there are not this case in most cases. For example, Schwarzschild spacetime (4) is a special solution of the Einstein's gravitational equations (3) in an assumption that all matters are spherically symmetric distributed in the universe with $T_{\mu\nu} = 0$ or vacuum. It is this kind of spacetimes that the standard model, the Big Bang hypothesis and black holes born on the universe. We are applying a special solution for characterizing the universe and believe it without a shadow of doubt in any place of the universe. However, there are infinite many solutions of Einstein's gravitational equations [7, 8]. But why the Schwarzschild spacetime was selected only for the universe because we are all fond of the symmetry and the uniformity on space, and we are firmly believing the spacetime structure of the universe should be so by observed datum of humans, at least in the nearby airspace of the earth.

Question II *Are the reality of things T really one of solutions of its equations*?

Science is established on an assumption that the reality or all behavior of a thing T can be characterized by mathematics, particularly, the second order differential equations in physics. However, the observation shows that a microcosmic particle is in two or more possible states of being, i.e., superposition such as the asking question of Schrödinger for the alive or dead of the cat in a box with poison switch. We can not even say which solution of Schrödinger equation (5) is the behavior of the particle because each solution is only one determined state in the eyes of humans.

Certainly, a reasonable or the multiverse interpretation on superposition of particles was presented by H. Everett in 1957. He explained the superposition of particles with an assumption that the wave function of an observer would be interacted with a superposed object [9] and concluded that different worlds in different quantum system obey equation (5) with an interpretation that the superposition of a particle

develops like a 2-branching universe. Thus, the answer of Question II is uncertain even if T is a microcosmic particle.

Today, it is just the Everett's multiverse interpretation on Schrödinger's cat enlightens humans known that the alive or dead of the cat is entangled and we can not say the cat is alive or dead separately. Philosophically, the Everett's multiverse notion on the superposition of particles is alluded in a famous fable, i.e., *the blind men with an elephant* or the formula (1). Today, this notion revolutionized changes an ambiguous interpretation that the reality of a thing T must be one but maybe all solutions of its differential equations and applies extensively to modern sciences. For example, it is the quark model that successfully classified all known elementary particles by mathematical symmetry but the quark model is indeed a multiverse and generally, all particles are nothing else but a multiverses [3] or complex networks in the microcosmic view.

Question III *Could the mathematics already characterizes the reality of things T?*

There is an exciting convincingness that mathematics can already characterizes the reality of all things, i.e., *Everything is Nothing Else but Mathematics* popularly in scientific community today, particularly, the *Mathematical Universe Hypothesis* in physics, a duplication of Pythagorean's assertion that "*Everything is a Number*". However, this notion is incorrect at least for today's mathematics because all mathematical systems should be homogenous without contradictions in logic. We can not conclude the equality

$$\text{Mathematical reality} \overset{\text{equal to}}{\longleftrightarrow} \text{Reality of things}$$

both in theory and practice. For instance, let H_1, H_2, H_3, H_4 and H_1', H_2', H_3', H_4' be two groups of horses constraint with running on respectively 4 straight lines

$$\textcircled{1}\begin{cases} x+y = 2 \\ x+y = -2 \\ x-y = -2 \\ x-y = 2 \end{cases} \quad \text{or} \quad \textcircled{2}\begin{cases} x = y \\ x+y = 4 \\ x = 2 \\ y = 2 \end{cases}$$

on the Euclidean plane $\mathbb{R}^2$. Clearly, the first system is non-solvable because $x+y=-2$ is contradictious to $x+y=2$, and so that for equations $x-y=-2$ and $x-y=2$ but the second system is solvable with $(x,y)=(2,2)$. *Could we conclude that the behavior of horses H_1', H_2', H_3', H_4' are a point $(2,2)$ and H_1, H_2, H_3, H_4 are nothing*? The answer is certainly not because all of the horses are running on the Euclidean plane $\mathbb{R}^2$ but we have known nothing by the solution of the two equation systems because the solvability of systems ① and ② only implies the orbits intersection in $\mathbb{R}^2$.

Why is this happening? It is because that while humans characterize a thing T in the universe by mathematics, it is usually complied with the compatible assumption of mathematics on T and often forgotten the original intention, i.e., hold on the reality of things T but have too much trust on the mathematical solution. Consequently, mathematics should be extended to include the non-mathematics for reality of things in the universe [1] because the contradictions exist everywhere in the eyes of humans. We can not conclude yet that mathematics can characterizes the reality of all things T in the universe until today.

4.2 Science's dilemma

Science's limitation naturally leads to a dilemma of science immediately. It gives the knowledge for humans but the knowledge is local or partial on things in the universe which always shows dual characters to humans, i.e., the beneficial or the harmful. However, it is easy to overstate the benefits but look without sees harms on a scientific achievement in a business community today. In this case, it is easy to breed the human's insatiable desires with immoderately abusing scientific achievements, and then brings a disaster finally to humans ourselves if it applies without constraints, particularly motivated only by the benefits of commercial interests. All of the harms come from the misunderstanding on science and incorrect applications of scientific achievements such as those cases following.

Physics has promoted the progress of humans but it also brought harmful things to human's living environment. For example, it pushes forward the aerospace industry which enables the exploration of humans on outer space. However, more and more satellites, space stations, probes, rocket debris and explosive fragments, working or abandoned are floating in space, disturbing the normally working of universe and also threatening the further exploration of humans because the aerocraft maybe collided with such an indefinable trashes in the space. Even in the daily life of humans we can also find the harms of applying physics. For example, the communicant equipments and facilities such as those of mobile phone, radio, TV station, microwave station bring convenience to humans but the radiate electromagnetic signals into space from time to time. However, it impacts on the health of humans, tested by the practice.

Chemistry has created new matters to meet the needs of humans but it caused complex problems simultaneously with its benefits to humans, for instance the environmental pollution, the resource depletion, the side effects of drugs, the pesticide residues and the lethal diseases such as cancer prevalence. Why these unpleasant things happen is because we have only a superficial understanding the fate, transport, toxicity on chemical products and without a comprehensive conclusion for their impacts on the environment and humans. For example, the plastics enables us protecting from wet but can not be degraded shortly by the nature, and we do not know the mechanism of accumulation of the pesticides in the food chain with impact on humans [10] until today.

Biology has brought benefits to humans but it presented

negative effects also to humans at the same time. For examples, the transgentic technology improved the resistance of crops but destroys those insects that feed on these pests which breaks the ecological chain, results in the structural change and deterioration of soil and the water pollution; the eating transgentic food maybe caused the modified gene invades human cells, produces pathogenic virus and harmful or lethal results including cancer and other negative effects; the cloning technology impacts on the nature and the social morality. Although the gene editing corrected or removed defect genes but will affects the normal functions of other cells at the same time, and while it reducing the genetic variation it maybe destroys a species just because of one disease, i.e., it increases the risk of infectious diseases and hypoimmunity or loss of other functions.

4.3 Out of dilemma

There are 2 sides for getting out of science's dilemma. One is the establishing of science in microcosmic level, i.e., *Microcosmic Science* for a complete understanding of things in the universe and two, is the self-awareness of human's ourself, the essence for out of the crisis.

4.3.1 Microcosmic science

As we know in the thinking pattern

$$\text{Matter} \overset{\overset{Decompose}{\longrightarrow}}{\underset{Abstract}{\longrightarrow}} \begin{matrix}\text{Microcosmic Particles}\\ \text{Complex Network}\end{matrix}$$

the reality of a thing T is the behavior with motivation of an abstracted complex network in the microcosmic level. Certainly, there are more microcosmic observing datum on the units, cells or microcosmic particles of matters by scientific instruments. Each of them appears in a space position at observing time and all of them are interrelated, for instance all cells in an animal. A microcosmic science is such a science established on the microcosmic datum of matters, including theory and experimental subjects. It must be established over 1-dimensional skeleton, i.e., topological graphs $\overrightarrow{G}$. However, we have no effective tools or methods, even no mathematics for such a work. Even though there is graph theory in mathematics but it is essentially discussing on binary relationship of elements without metrics, can not be immediately applied to understand the reality of matters, particularly, the microcosmic science.

Could we establish such a mathematics over topological graphs for microcosmic science? The answer is positive inspired by traditional Chinese medicine. Certainly, there are 12 meridians which completely reflects the physical condition of human body in traditional Chinese medicine [11], i.e., the lung meridian of hand-TaiYin (LU), the large intestine meridian of hand YangMing (LI), the stomach meridian of foot-YangMing (ST), the spleen meridian of foot-TaiYin (SP), the heart meridian of hand-ShaoYin (HT), the small intestine meridian of hand-TaiYang (SI), the urinary blandder meridian of foot-TaiYang (BL), the kidney meridian of foot-ShaoYin (KI), the pericardium meridian of hand-JueYin (PC), the sanjiao meridian of hand-ShaoYang (SJ), the gall bladder meridian of foot-ShaoYang (GB), the liver meridian of foot-JueYin (LR) in *Standard China National Standard* (GB 12346-90).

Notice that maintaining the balance of Yin (Y^-) with that of Yang (Y^+) is the foundation of Chinese culture, particularly on a healthy human body. According to the view of traditional Chinese medicine, if there exists an imbalanced acupoint on one of the 12 meridians this person must has illness and in turn, there must be imbalance acupoints on the 12 meridians for a patient. Thus, finding out which acupoint on which meridian is imbalance with Y^- more than Y^+ or Y^+ more than Y^- is the main duty of a Chinese doctor. Then, by the natural ruler, i.e., *reducing the excess with supply the insufficient* of the universe, the doctor regulates the meridian by acupuncture or drugs so that the patient recovers balance on the imbalance acupoint [11], which is the essential treatment of traditional Chinese medicine.

Although a matter can be infinitely subdivided into submatters, the success of traditional Chinese medicine implies that there exists an inherited a topological skeleton or graph G in things, particularly, human body in the universe. By view of biology, there are only 2 kinds of things, i.e., living or death body which suggest 2 mathematical elements holding with conservation laws for things in the universe in the microcosmic level following:

Element 1 (Non-Living Body). A *continuity flow* $\overrightarrow{G}^L$ is an oriented embedded graph $\overrightarrow{G}$ in a topological space $\mathcal{S}$ associated with a mapping $L : v \to L(v)$, $(v,u) \to L(v,u)$, 2 end-operators $A^+_{vu} : L(v,u) \to L^{A^+_{vu}}(v,u)$ and $A^+_{uv} : L(u,v) \to L^{A^+_{uv}}(u,v)$ on a Banach space $\mathfrak{B}$ over a field $\mathfrak{F}$ with $L(v,u) = -L(u,v)$, $A^+_{vu}(-L(v,u)) = -L^{A^+_{vu}}(v,u)$ for $\forall (v,u) \in E\left(\overrightarrow{G}\right)$ and holding with continuity equation

$$\sum_{u\in N_G(v)} L^{A^+_{vu}}(v,u) = L(v) \ \text{ for } \ \forall v \in V\left(\overrightarrow{G}\right).$$

Element 2 (Living Body). A *harmonic flow* $\overrightarrow{G}^L$ is an oriented embedded graph $\overrightarrow{G}$ in a topological space $\mathcal{S}$ associated with a mapping $L : v \to L(v) - iL(v)$ for $v \in E\left(\overrightarrow{G}\right)$ and $L : (v,u) \to L(v,u) - iL(v,u)$, 2 end-operators $A^+_{vu} : L(v,u) - iL(v,u) \to L^{A^+_{vu}}(v,u) - iL^{A^+_{vu}}(v,u)$ and $A^+_{uv} : L(v,u) - iL(v,u) \to L^{A^+_{uv}}(v,u) - iL^{A^+_{uv}}(v,u)$ on a Banach space $\mathfrak{B}$ over a field $\mathfrak{F}$, where $i^2 = -1$, $L(v,u) = -L(u,v)$ for $\forall (v,u) \in E\left(\overrightarrow{G}\right)$ and

holding with continuity equation

$$\sum_{u\in N_G(v)}\left(L^{A^+_{vu}}(v,u)-iL^{A^+_{vu}}(v,u)\right)=L(v)-iL(v)$$

for $\forall v\in V\left(\overrightarrow{G}\right)$.

Notice that if we let the Banach space to be $\mathfrak{B}\times\mathfrak{B}$ then the Element 2 is only a special Element 1 with complex vector. However, it reflects living bodies with respective real, imaginary parts $L(v,u)$, $-L(v,u)$ appearing in pair. If one lost then the counterpart is no longer exists, i.e., it is depth. This notion can be also used to explain the entangled state, i.e., the alive or dead of Schrödinger's cat in the box by a complex state $A-iA$ in such a way that alive for an $A\neq\mathbf{0}$ but dead if $A=\mathbf{0}$.

According to the structure of the 12 meridians on human body, we can classify them into 3 classes, i.e., *paths*: LU, LI, SP, HT, SI, KI, PC, LR; *trees*: GB, ST, SJ and a circuit attached with 2 pathes P_{m_1}, P_{m_2}: BL. Define an oriented graph

$$\begin{aligned}\overrightarrow{G} \;=\;& P_{11}(LU)\bigcup P_{20}(LI)\bigcup P_{21}(SP)\bigcup\\ & P_9(HT)\bigcup P_{19}(SI)\bigcup P_{27}(KI)\bigcup\\ & P_9(PC)\bigcup P_{14}(LR)\bigcup T_{44}(GB)\bigcup\\ & T_{45}(ST)\bigcup T_{23}(SJ)\bigcup G_{67}(BL)\bigcup\\ & P_{28}(DU)\bigcup P_{24}(RN)\end{aligned}$$

with orientations:

chest→hand→head→foot→chest

in human body and $L:\ v\in V(\overrightarrow{G})\to L(v)-iL(v)$ and $L:\ (v,u)\in E(\overrightarrow{G})\to L(v,u)-iL(v,u)$, where DU and RN are respectively the DU and REN meridians on human body, $P_n(X), T_n(X)$ and $G_n(X)$ denote the path, tree or graph of meridian of order n. Then, $\overrightarrow{G}^L$ is nothing else but a harmonic flow equivalent to human body by the view of traditional Chinese medicine, a kind of Element 2 of order 361.

As shown in references [7, 8, 12, 13], the Elements 1 and 2 can be applied to characterize the behavior of things T in the universe with $\overrightarrow{G}^L$ a globally mathematical elements in the sense that if $\mathfrak{G}$ is a closed family of graphs under union operation, $\mathfrak{B}$ is a Banach or Hilbert space, then all Elements 1 or 2, i.e., $\overrightarrow{G}^L$ with $\overrightarrow{G}\in\mathfrak{G}$ respectively form a Banach or Hilbert flow space and closed under the action of differential and integral operators with a few generalized theorems in functionals. Particularly, they can be used to characterize the dynamic behavior of things T, living or non-living bodies in the universe by Euler-Lagrange equations

$$\frac{\partial\overrightarrow{G}^L}{\partial x_i}-\frac{d}{dt}\frac{\partial\overrightarrow{G}^L}{\partial\dot{x}_i}=\mathbf{O},\quad 1\leq i\leq n$$

where, $\overrightarrow{G}^L$ is the harmonic or continuity flow inherited in T, $L(v,u)$ is the Lagrangian on edge (v,u) and $\mathbf{O}$ is the zero-flow $\overrightarrow{G}^{\mathbf{0}}$, i.e., a labeling $\mathbf{0}:(v,u)\to\mathbf{0}$ for $(v,u)\in E\left(\overrightarrow{G}\right)$.

5 Human self-awareness

The original intention of science is to understand things in the universe, promote the survival and development of humans ourself and then, construct a harmonious system of humans with the nature. Historically, human's experience verified times that the more intruding with higher damage of humans on the nature, the more serious nature's punishments on humans society are. The leader is nothing else but humans ourself in the couple of humans and the nature. As discussed in the previous. Science has itself limitation and all of its achievements is only the local or the partial, and what humans gotten maybe always a local conclusion on the reality of things in the universe. For example, humans have not really understood the internal and external mechanism of planets, only hold on their's laws by observations. In this case, discussing the capture of asteroids for energy or human alien migration is not realistic, and the result in harming to the universe is immeasurable.

Hence, science needs returning to a rational research on the respecting with protecting of the nature, and abandons the idea that humans are the center and wish to govern the universe by a limited understanding. Furthermore, we are need also to distinguish a scientific research is for human survival with development or only serves to human's enjoying because the later is causing the loss of human's natural instincts sometimes. Science should returns to the theme of harmonious development of humans with the nature. While researching a scientific problem, it should takes more times on the maybe harming to humans and to the universe with extents for its application. In this case, *is to discuss the destruction of the earth then migrates to other planets or develops with the earth*? In addition, *is to research the destruction of our universe and then migrates to other universes*? The answer is obvious because we have only one earth and one universe on which we live. It can be only harmonious with but not destroying the earth or the universe if we would like to a sustainable developing. Even if it were necessary to exploit the resources of the earth or the universe, we should also be minimized the natural intrusion and maximized the use of natural resources constraint with a model of circular development.

We have faced survival problems such as those of population growth, food safety, health, environmental pollution and resource depletion today. However, the greatest crisis facing humans is not the poverty or unfair allocation of natural resources but the greed with ignorance, and hopefully to govern the universe by our own understanding or a realization dependent on local or partial perception of the nature, particularly, the abusing of scientific achievements such as those of the overdevelop or overuse of resources, vehicles, internet,

farm chemicals and biological products. The main step for out of the crisis needs the human self-awareness, i.e., abandoning their arrogance and developing harmoniously with the nature because we have only a local or partial understanding for reality of things in the universe. Even though we have established science on things T, it is only an understanding of humans ourself on the earth, maybe not the reality of things in the whole universe. Thus, the only viable way for human's continually generations is to develop with the nature.

Submitted on May 7, 2019

References

1. Mao L. Mathematics on non-mathematics - A combinatorial contribution. *International J. Math. Combin.*, 2014, v. 3, 1–34.
2. Smarandache F. *Paradoxist Geometry*, State Archives from Valcea, Rm. Valcea, Romania, 1969, and in *Paradoxist Mathematics*, Collected Papers (Vol. II), Kishinev University Press, Kishinev, 5–28, 1997.
3. Tegmark M. Parallel universes, in *Science and Ultimate Reality: From Quantum to Cosmos*, ed. by J.D. Barrow, P.C.W. Davies and C.L. Harper, Cambridge University Press, 2003.
4. Mao L. Combinatorial Geometry with Applications to Field Theory. The Education Publisher Inc., USA, 2011.
5. Nambu Y. Quarks: Frontiers in Elementary Particle Physics. World Scientific Publishing Co. Pte. Ltd, 1985.
6. Smarandache F. and Rabounski D. Unmatter entities inside nuclei, predicted by the Brightsen nucleon cluster model. *Progress in Physics*, 2006, no. 1, 14–18.
7. Mao L. Extended Banach $\overrightarrow{G}$-flow spaces on differential equations with applications, *Electronic J. Mathematical Analysis and Applications*, 2015, v. 3, no. 2, 59–91.
8. Mao L. A review on natural reality with physical equation. *Progress in Physics*, 2015, v. 11, 276–282.
9. Everett H. Relative state formulation of quantum mechanics. *Rev. Mod. Phys.*, 1957, v. 29, 454–462.
10. Committee on Mathematical Sciences Research for DOE's Computational Biology and National Research Council of the National Academies, *Mathematics and 21st Century Biology*, National Academy Press, USA, 2005.
11. Zhang Z. Comments on the Inner Canon of Emperor (Qing Dynasty, in Chinese), Northern Literature and Art Publishing House, 2007.
12. Mao L. Complex system with flows and synchronization. *Bull. Cal. Math. Soc.*, 2017, v. 109, no. 6, 461–484.
13. Mao L. Harmonic flow's dynamics on animals in microscopic level with balance recovery. *International J. Math. Combin.*, 2019, v. 1, 1–44.

The Origin of Inertial Mass in the Spacetime Continuum

Pierre A. Millette

E-mail: PierreAMillette@alumni.uottawa.ca, Ottawa, Canada

In this paper, we revisit the nature of inertial mass as provided by the Elastodynamics of the Spacetime Continuum (*STCED*). We note that, in addition to providing a physical explanation for inertial mass and for wave-particle duality, it answers unresolved questions pertaining to mass: It provides a direct physical definition of mass independent of the operational definition of mass currently used. It shows that, in general, a singular "point" particle is not physically valid and that particles need to be given a finite volume to avoid invalid results and give physically realistic ones. It confirms theoretically the equivalence of inertial and gravitational mass. It demonstrates that Mach's principle (or conjecture) is incorrect in that inertia originates from the massive dilatation associated with a spacetime deformation, not from interaction with the average mass of the universe. It shows that the electromagnetic field is transverse and massless, and that it contributes to the particle's total energy, but not to its inertial mass.

It must also be said that the origin of inertia is and remains ***the*** *most obscure subject in the theory of particles and fields. A. Pais, 1982 [1, p. 288]*

... the notion of mass, although fundamental to physics, is still shrouded in mystery. M. Jammer, 2000 [2, p. ix]

1 Introduction

In this paper, we revisit the nature of inertial mass as provided by the Elastodynamics of the Spacetime Continuum (*STCED*) [3, 4]. *STCED* is a natural extension of Einstein's General Theory of Relativity which blends continuum mechanical and general relativistic descriptions of the spacetime continuum. The introduction of strains in the spacetime continuum as a result of the energy-momentum stress tensor allows us to use, by analogy, results from continuum mechanics, in particular the stress-strain relation, to provide a better understanding of the general relativistic spacetime.

2 Elastodynamics of the Spacetime Continuum

The stress-strain relation for an isotropic and homogeneous spacetime continuum is given by [3, 4]

$$2\bar{\mu}_0\,\varepsilon^{\mu\nu} + \bar{\lambda}_0\,g^{\mu\nu}\varepsilon = T^{\mu\nu} \tag{1}$$

where $\bar{\lambda}_0$ and $\bar{\mu}_0$ are the Lamé elastic constants of the spacetime continuum: $\bar{\mu}_0$ is the shear modulus (the resistance of the spacetime continuum to *distortions*) and $\bar{\lambda}_0$ is expressed in terms of $\bar{\kappa}_0$, the bulk modulus (the resistance of the spacetime continuum to *dilatations*):

$$\bar{\lambda}_0 = \bar{\kappa}_0 - \bar{\mu}_0/2 \tag{2}$$

in a four-dimensional continuum. $T^{\mu\nu}$ is the general relativistic energy-momentum stress tensor, $\varepsilon^{\mu\nu}$ the spacetime continuum strain tensor resulting from the stresses, and

$$\varepsilon = \varepsilon^{\alpha}{}_{\alpha}\,, \tag{3}$$

the trace of the strain tensor obtained by contraction, is the volume dilatation ε defined as the change in volume per original volume [9, see pp. 149–152] and is an invariant of the strain tensor. It should be noted that the structure of (1) is similar to that of the field equations of general relativity,

$$R^{\mu\nu} - \tfrac{1}{2}\,g^{\mu\nu}R = -\varkappa\,T^{\mu\nu} \tag{4}$$

where $R^{\mu\nu}$ is the Ricci curvature tensor, R is its trace, $\varkappa = 8\pi G/c^4$ and G is the gravitational constant (see [3, Ch. 2] for more details).

3 Inertial mass in *STCED*

In *STCED*, as shown in [3,4], energy propagates in the spacetime continuum (*STC*) as wave-like deformations which can be decomposed into *dilatations* and *distortions*. *Dilatations* involve an invariant change in volume of the spacetime continuum which is the source of the associated rest-mass energy density of the deformation. On the other hand, *distortions* correspond to a change of shape (shearing) of the spacetime continuum without a change in volume and are thus massless.

Thus deformations propagate in the spacetime continuum by longitudinal (*dilatation*) and transverse (*distortion*) wave displacements. This provides a natural explanation for wave-particle duality, with the massless transverse mode corresponding to the wave aspects of the deformations and the massive longitudinal mode corresponding to the particle aspects of the deformations.

The rest-mass energy density of the longitudinal mode is given by [4, see Eq. (32)]

$$\rho c^2 = 4\bar{\kappa}_0\varepsilon \tag{5}$$

where ρ is the rest-mass density, c is the speed of light, $\bar{\kappa}_0$ is the bulk modulus of the *STC*, and ε is the volume dilatation given by (3). Integrating over the 3-D space volume,

$$\int_{V_3} \rho c^2\,\mathrm{d}V_3 = 4\bar{\kappa}_0 \int_{V_3} \varepsilon\,\mathrm{d}V_3\,, \tag{6}$$

and using

$$m = \int_{V_3} \rho \, \mathrm{d}V_3 \tag{7}$$

in (6), where m is the rest mass (often denoted as m_0) of the deformation, we obtain

$$mc^2 = 4\bar{\kappa}_0 \, V_{\varepsilon s} \tag{8}$$

where

$$V_{\varepsilon s} = \int_{V_3} \varepsilon \, \mathrm{d}V_3 \tag{9}$$

is the space volume dilatation corresponding to rest-mass m, and spacetime continuum volume dilatation ε is the solution of the 4-D dilatational (longitudinal) wave equation [3, see Eq. (3.35)]

$$(2\bar{\mu}_0 + \bar{\lambda}_0) \, \nabla^2 \varepsilon = -\partial_\nu X^\nu \tag{10}$$

where ∇ and ∂ are the 4-D operators and X^ν is the spacetime continuum volume force.

This demonstrates that mass is not independent of the spacetime continuum, but rather mass is part of the spacetime continuum fabric itself. Hence mass results from the dilatation of the spacetime continuum in the longitudinal propagation of energy-momentum in the spacetime continuum. Matter does not warp spacetime, but rather, matter *is* warped spacetime (*i.e.* dilated spacetime). The universe consists of the spacetime continuum and energy-momentum that propagates in it by deformation of its structure.

It is interesting to note that Pais, in his scientific biography of Einstein *'Subtle is the Lord...'*, mentions [1, p. 253]

> The trace of the energy momentum tensor does vanish for electromagnetic fields but not for matter.

which is correct, as shown in [5, 6], where the zero trace of the electromagnetic field energy-momentum stress tensor is reflected in the zero mass of the photon. The missing link in general relativity is the understanding that the trace of the energy-momentum stress tensor is related to the trace of the spacetime continuum strain tensor and is proportional to the mass of matter as given by (5) and (8).

There are basic questions of physics that can be resolved given this understanding of the origin of inertial mass. The following sections deal with many of these unresolved questions.

3.1 Definition of mass

An important consequence of relations (5) and (8) is that they provide a definition of mass. The definition of mass is still one of the open questions in physics, with most authors adopting an indirect definition of mass based on the ratio of force to acceleration [15, see Ch. 8]. However, mass is one of the fundamental dimensions of modern systems of units, and as such, should be defined directly, not indirectly. This is a reflection of the current incomplete understanding of the nature of mass in physics. *STCED* provides a direct physical definition of mass: *mass is the invariant change in volume of spacetime in the longitudinal propagation of energy-momentum in the spacetime continuum*.

Note that the operational definition of mass ($m = F/a$) is still needed to measure the mass of objects and compare them. Jammer covers the various operational and philosophical definitions of mass that have been proposed [2, Ch. 1].

3.2 Point particles

The fact that the mass of a particle corresponds to a finite spacetime volume dilatation $V_{\varepsilon s}$ shows that a singular "point" particle is not physically valid. All particles occupy a finite volume, even if that volume can be very small. Problems arising from point particles are thus seen to result from the abstraction of representing some particles as point objects. Instead, particles need to be given a finite volume to give physically realistic results and avoid invalid results.

3.3 Equivalence of inertial and gravitational mass

Einstein's general relativistic principle of equivalence of inertial and gravitational mass can be given added confirmation in *STCED*. As shown in [5, 7], the Ricci tensor can also be decomposed into dilatation and distortion components. The dilatation component can be shown to result in Poisson's equation for a newtonian gravitational potential [3, see Eq. (2.44)] where the gravitational mass density is identical to the rest-mass density identified in *STCED*. This confirms theoretically the equivalence of inertial mass and gravitational mass, as demonstrated experimentally within the accuracy currently achievable [10].

3.4 Mach's principle

Mach's principle, a terminology first used by Einstein [1, p. 287], was not explicitly stated by Mach, and hence various takes on its statement exist. One of the better formulation holds that one can determine rotation and hence define inertial frames with respect to the fixed stars [11, see pp. 86–88]. By extension, inertia would then be due to an interaction with the average mass of the universe [11, see p. 17].

This principle played an important role in the initial development of general relativity by Einstein which is well documented by Pais [1, pp. 283–287]. It also had an impact on the initial work performed in cosmology by Einstein who was searching for a cosmological model that would be in accord with Mach's principle. Einstein's evolving perspective on Mach's work is best summarized by Pais [1, p. 287]:

> So strongly did Einstein believe at that time in the relativity of inertia that in 1918 he stated as being on equal footing three principles on which a satisfactory theory of gravitation should rest [Mach's principle was the third] ... In later years, Einstein's enthusiasm for Mach's principle waned and finally vanished.

Modifications of Einstein's Theory of General Relativity have been proposed in an attempt to incorporate Mach's principle into general relativity (see for example [12, 13]).

The book *Gravitation and Inertia* by Ciufolini and Wheeler [14], with its emphasis on geometrodynamics and its well-known sayings "spacetime tells mass how to move and mass tells spacetime how to curve" and "inertia here arises from mass there", explores these ideas in detail. However, it is important to realize that this perspective is an interpretation of Einstein's field equations of general relativity (4). These equations are simply a relation between the geometry of the spacetime continuum and the energy-momentum present in its structure. *STCED* shows that mass is not outside of the spacetime continuum telling it how to curve (so to speak), but rather mass is part of the spacetime continuum fabric itself participating in the curvature of the spacetime continuum. The geometry of the spacetime continuum is generated by the combination of all spacetime continuum deformations which are composed of longitudinal massive dilatations and transverse massless distortions.

As shown in [3, §2.5], the geometry of spacetime used in (4) can thus be considered to be a linear composition (represented by a sum) of *STC* deformations, starting with the total energy-momentum generating the geometry of general relativity, $T^{\mu\nu}_{GR}$, being a composition of the energy-momentum of the individual deformations of *STCED*, $T^{\mu\nu}_{STCED}$:

$$T^{\mu\nu}_{GR} = \sum T^{\mu\nu}_{STCED} \,. \tag{11}$$

Substituting into (11) from (1) and (4), we obtain

$$-\frac{1}{\varkappa}\left[R^{\mu\nu} - \tfrac{1}{2}\, g^{\mu\nu} R\right] = \sum \left[2\bar{\mu}_0\, \varepsilon^{\mu\nu} + \bar{\lambda}_0\, g^{\mu\nu}\varepsilon\right] . \tag{12}$$

Contraction of (12) yields the relation

$$\frac{1}{\varkappa}\, R = \sum 2(\bar{\mu}_0 + 2\bar{\lambda}_0)\, \varepsilon \tag{13}$$

which, using (2) and (5), simplifies to

$$\frac{1}{\varkappa}\, R = \sum 4\, \bar{\kappa}_0\, \varepsilon = \sum \rho\, c^2 \tag{14}$$

i.e. the curvature of the spacetime continuum arises from the composition of the effect of individual deformations and is proportional to the rest-mass energy density present in the spacetime continuum. Substituting for $R/\varkappa$ from (14) into (12), and rearranging terms, we obtain

$$\frac{1}{\varkappa}\, R^{\mu\nu} = \sum \left[(\bar{\lambda}_0 + \bar{\mu}_0) g^{\mu\nu}\varepsilon - 2\bar{\mu}_0\, \varepsilon^{\mu\nu}\right] . \tag{15}$$

Eqs. (14) and (15) give the relation between the microscopic description of the strains (*i.e.* deformations of the spacetime continuum) and the macroscopic description of the gravitational field in terms of the curvature of the spacetime continuum resulting from the combination of the many microscopic displacements of the spacetime continuum from equilibrium. The source of the inertia is thus in the massive dilatation associated with each deformation, and Mach's principle (or conjecture as it is also known) is seen to be incorrect.

3.5 Electromagnetic mass

The advent of Maxwell's theory of electromagnetism in the second half of the nineteenth century led to the possibility of inertia resulting from electromagnetism, first proposed in 1881 by J. J. Thomson [15, see Chapter 11]. The application of the concept of electromagnetic mass to the electron discovered by J. J. Thomson in 1897, by modelling it as a small charged sphere, led to promising results [16, see Chapter 28]. One can then calculate the energy in the electron's electric field and divide the result by c^2. Alternatively, the electromagnetic momentum of a moving electron can be calculated from Poynting's vector and the electromagnetic mass set equal to the factor multiplying the electron's velocity vector. Different methods give different results.

Using the classical electron radius

$$r_0 = \frac{e^2}{m_e c^2} \tag{16}$$

where e is the electronic charge and m_e the mass of the electron, then the electromagnetic mass of the electron can be written as

$$m_{em} = k_e\, \frac{e^2}{r_0 c^2} \tag{17}$$

where the factor k_e depends on the assumed charge distribution in the sphere and the method of calculation used. For a surface charge distribution, $k_e = 2/3$, while for a uniform volume distribution, $k_e = 4/5$. Numerous modifications were attempted to get $m_{em} = m_e$ [15, 16] with Poincaré introducing non-electrical forces known as "Poincaré stresses" to get the desired result. This is a classical treatment that does not take relativistic or quantum effects into consideration.

It should be noted that the simpler classical treatment of the electromagnetic mass of the electron based purely on the electric charge density of the electron is a calculation of the static mass of the electron. In *STCED*, the charge density ϱ can be calculated from the current density four-vector j^ν (see [3, §4.3])

$$j^\nu = \frac{\varphi_0}{\mu_0}\, \frac{2\bar{\mu}_0 + \bar{\lambda}_0}{2\bar{\mu}_0}\, \varepsilon^{;\nu} \tag{18}$$

where φ_0 is the *STC electromagnetic shearing potential constant*, which has units of $[\mathrm{V}\cdot\mathrm{s}\cdot\mathrm{m}^{-2}]$ or equivalently [T], μ_0 is the electromagnetic permeability of free space, and $\varepsilon^{;\nu}$ can be written as the dilatation current $\xi^\nu = \varepsilon^{;\nu}$. Substituting for j^ν from (18) in the relation [23, see p. 94]

$$j^\nu j_\nu = \varrho^2 c^2, \tag{19}$$

we obtain the expression for the charge density

$$\varrho = \frac{1}{2}\,\frac{\varphi_0}{\mu_0 c}\,\frac{2\bar{\mu}_0 + \bar{\lambda}_0}{2\bar{\mu}_0}\,\sqrt{\varepsilon^{;\nu}\varepsilon_{;\nu}}\,. \tag{20}$$

Note the difference between the electromagnetic permeability of free space μ_0 and the Lamé elastic constant $\bar{\mu}_0$ used to denote the spacetime continuum shear modulus.

We see that the charge density derives from the norm of the gradient of the volume dilatation ε, *i.e.*

$$\begin{aligned}\|\varepsilon^{;\nu}\| &= \sqrt{\varepsilon^{;\nu}\varepsilon_{;\nu}} \\ &= \sqrt{\left(\frac{\partial\varepsilon}{\partial x}\right)^2 + \left(\frac{\partial\varepsilon}{\partial y}\right)^2 + \left(\frac{\partial\varepsilon}{\partial z}\right)^2 + \frac{1}{c^2}\left(\frac{\partial\varepsilon}{\partial t}\right)^2}\end{aligned} \tag{21}$$

in cartesian coordinates, and from the above, (20) becomes

$$\varrho = \frac{1}{2}\,\frac{\varphi_0}{\mu_0 c}\,\frac{2\bar{\mu}_0 + \bar{\lambda}_0}{2\bar{\mu}_0}\,\|\varepsilon^{;\nu}\|\,. \tag{22}$$

The charge density is a manifestation of the spacetime fabric itself, however it does not depend on the volume dilatation ε, only on its gradient, and it does not contribute to inertial mass as given by (5). The electromagnetic mass calculation is based on the energy in the electron's electric field and we now consider electromagnetic field energy in *STCED* to clarify its contribution, if any, to inertial mass. This also covers the calculation of electromagnetic mass from the Poynting vector.

3.6 Electromagnetic field energy in the spacetime continuum

As shown in [8], the correct special relativistic relation for momentum p is given by

$$p = m_0\, u\,, \tag{23}$$

where m_0 is the proper or rest mass, u is the velocity with respect to the proper time τ, given by $u = \gamma v$, where

$$\gamma = \frac{1}{(1-\beta^2)^{1/2}}\,, \tag{24}$$

$\beta = v/c$, and v is the velocity with respect to the local time t. When dealing with dynamic equations in the local time t instead of the invariant proper time τ, momentum p is given by

$$p = m^* v\,, \tag{25}$$

where the relativistic mass m^* is given by

$$m^* = \gamma m_0\,. \tag{26}$$

Eq. (25), compared to (23), shows that relativistic mass m^* is an effective mass which results from dealing with dynamic equations in the local time t instead of the invariant proper time τ. The relativistic mass energy m^*c^2 corresponds to the total energy of an object (invariant proper mass plus kinetic energy) measured with respect to a given frame of reference [8]. As noted by Jammer [2, p. 41],

> Since [velocity v] depends on the choice of [reference frame] S relative to which it is being measured, [relativistic mass m^*] also depends on S and is consequently a relativistic quantity and not an intrinsic property of the particle.

Using the effective mass, we can write the energy E as the sum of the proper mass and the kinetic energy K of the body, which is typically written as

$$E = m^* c^2 = m_0\, c^2 + K\,. \tag{27}$$

If the particles are subjected to forces, these stresses must be included in the energy-momentum stress tensor, and hence added to K. Thus we see that the inertial mass corresponds to the proper or rest mass of a body, while relativistic mass does not represent an actual increase in the inertial mass of a body, just its total energy (see Taylor and Wheeler [17], Okun [18–20], Oas [21, 22]).

Considering the energy-momentum stress tensor of the electromagnetic field, we can show that $T^{\alpha}{}_{\alpha} = 0$ as expected for massless photons, while

$$T^{00} = \frac{\epsilon_0}{2}\,(E^2 + c^2 B^2) = U_{em} \tag{28}$$

is the total energy density, where U_{em} is the electromagnetic field energy density, ϵ_0 is the electromagnetic permittivity of free space, and E and B have their usual significance for the electric and magnetic fields (see [3, §5.3]). As $m_0 = 0$ for the electromagnetic field, the electromagnetic field energy then needs to be included in the K term in (27).

In general, the energy relation in special relativity is quadratic, given by

$$E^2 = m_0^2\, c^4 + p^2 c^2\,, \tag{29}$$

where p is the momentum. Making use of the effective mass (26) allows us to obtain (25) from (29) [6], starting from

$$m^{*2} c^4 = \gamma^2 m_0^2\, c^4 = m_0^2\, c^4 + p^2 c^2\,. \tag{30}$$

This section provides a description of the electromagnetic field energy using a quadratic energy relation which corresponds to the more complete classical treatment of the electromagnetic mass of the electron based on the Poynting vector of the electron in motion.

In *STCED*, energy is stored in the spacetime continuum as strain energy [6]. As seen in [4, see Section 8.1], the strain energy density of the spacetime continuum is separated into two terms: the first one expresses the dilatation energy density (the mass longitudinal term) while the second one expresses the distortion energy density (the massless transverse term):

$$\mathcal{E} = \mathcal{E}_{\parallel} + \mathcal{E}_{\perp} \tag{31}$$

where

$$\mathcal{E}_{\parallel} = \frac{1}{2}\,\bar{\kappa}_0 \varepsilon^2 \equiv \frac{1}{32\bar{\kappa}_0}\,\rho^2 c^4\,, \tag{32}$$

ρ is the rest-mass density of the deformation, and

$$\mathcal{E}_\perp = \bar{\mu}_0\, e^{\alpha\beta} e_{\alpha\beta} = \frac{1}{4\bar{\mu}_0}\, t^{\alpha\beta} t_{\alpha\beta}\,, \tag{33}$$

with the strain distortion

$$e^{\alpha\beta} = \varepsilon^{\alpha\beta} - e_s\, g^{\alpha\beta} \tag{34}$$

and the strain dilatation $e_s = \frac{1}{4}\,\varepsilon^{\alpha}{}_{\alpha}$. Similarly for the stress distortion $t^{\alpha\beta}$ and the stress dilatation t_s. Then the dilatation (massive) strain energy density of the deformation is given by the longitudinal strain energy density (32) and the distortion (massless) strain energy density of the deformation is given by the transverse strain energy density (33).

As shown in [3, §5.3.1] for the electromagnetic field, the longitudinal term is given by

$$\mathcal{E}_\parallel = 0 \tag{35}$$

as expected [24, see pp. 64–66]. This result thus shows that the rest-mass energy density of the electromagnetic field, and hence of the photon is zero, *i.e.* the photon is massless. The transverse term is given by [3, §5.3.2]

$$\mathcal{E}_\perp = \frac{1}{4\bar{\mu}_0}\left[\epsilon_0^2\left(E^2 + c^2B^2\right)^2 - \frac{4}{c^2}\,S^2\right] \tag{36}$$

or

$$\mathcal{E}_\perp = \frac{1}{\bar{\mu}_0}\left[{U_{em}}^2 - \frac{1}{c^2}\,S^2\right] \tag{37}$$

where $U_{em} = \frac{1}{2}\,\epsilon_0(E^2 + c^2B^2)$ is the electromagnetic field energy density as before and S is the magnitude of the Poynting vector. The Poynting four-vector is defined as [3, §5.4]

$$S^\nu = (cU_{em}, \boldsymbol{S})\,, \tag{38}$$

where U_{em} is the electromagnetic field energy density, and $\boldsymbol{S}$ is the Poynting vector. Furthermore, S^ν satisfies

$$\partial_\nu S^\nu = 0\,. \tag{39}$$

Using definition (38) in (37), we obtain the transverse massless energy density of the electromagnetic field

$$\mathcal{E}_\perp = \frac{1}{\bar{\mu}_0 c^2}\, S_\nu S^\nu\,. \tag{40}$$

The indefiniteness of the location of the field energy referred to by Feynman [16, see p. 27-6] is thus resolved: the electromagnetic field energy resides in the distortions (transverse displacements) of the spacetime continuum.

Hence the electromagnetic field is transverse and massless, and has no massive longitudinal component. The electromagnetic field has energy, but no rest mass, and hence no inertia. From *STCED*, we see that electromagnetism as the source of inertia is not valid.

Electromagnetic mass is thus seen to be an unsuccessful attempt to account for the inertial mass of a particle from its electromagnetic field energy. The electromagnetic field contributes to the particle's total energy, but not to its inertial mass which *STCED* shows originates in the particle's dilatation energy density (the mass longitudinal term) which is zero for the electromagnetic field.

4 Discussion and conclusion

In this paper, we have revisited the nature of inertial mass as provided by the Elastodynamics of the Spacetime Continuum (*STCED*) which provides a better understanding of general relativistic spacetime. Mass is shown to be the invariant change in volume of spacetime in the longitudinal propagation of energy-momentum in the spacetime continuum. Hence mass is not independent of the spacetime continuum, but rather mass is part of the spacetime continuum fabric itself.

STCED provides a direct physical definition of mass. In addition, it answers many of the unresolved questions that pertain to the nature of mass:

- The mass of a particle corresponds to a finite spacetime volume dilatation $V_{\varepsilon s}$ and particles need to be given a finite volume (as opposed to "point particles") to give physically realistic results and avoid invalid results.
- It confirms theoretically the equivalence of inertial and gravitational mass.
- The source of inertia is in the massive dilatation associated with each deformation, and Mach's principle (or conjecture), which holds that inertia results from interaction with the average mass of the universe, is seen to be incorrect.
- The electromagnetic field is transverse and massless, and has no massive longitudinal component. It has energy, but no rest mass, and hence no inertia. The electromagnetic field contributes to the particle's total energy, but not to its inertial mass.

STCED thus provides a physical model of the nature of inertial mass, which also includes an explanation for wave-particle duality. This model leads to the clarification and resolution of unresolved and contentious questions pertaining to inertial mass and its nature.

Received on May 27, 2019

References

1. Pais A. 'Subtle is the Lord...': The Science and the Life of Albert Einstein. Oxford University Press, Oxford, 1982.
2. Jammer M. Concepts of Mass in Contemporary Physics and Philosophy. Princeton University Press, Princeton, 2000.
3. Millette P. A. Elastodynamics of the Spacetime Continuum: A Spacetime Physics Theory of Gravitation, Electromagnetism and Quantum Physics. American Research Press, Rehoboth, NM, 2017.

4. Millette P. A. Elastodynamics of the Spacetime Continuum. *The Abraham Zelmanov Journal*, 2012, vol. 5, 221–277.
5. Millette P. A. On the Decomposition of the Spacetime Metric Tensor and of Tensor Fields in Strained Spacetime. *Progress in Physics*, 2012, vol. 8 (4), 5–8.
6. Millette P. A. Strain Energy Density in the Elastodynamics of the Spacetime Continuum and the Electromagnetic Field. *Progress in Physics*, 2013, vol. 9 (2), 82–86.
7. Millette P. A. Dilatation–Distortion Decomposition of the Ricci Tensor. *Progress in Physics*, 2013, vol. 9 (4), 32–33.
8. Millette P. A. On Time Dilation, Space Contraction, and the Question of Relativistic Mass. *Progress in Physics*, 2017, vol. 13 (4), 202–255.
9. Segel L. A. Mathematics Applied to Continuum Mechanics. Dover Publications, New York, 1987.
10. Roll P. G., Krotkov R., and Dicke R. H. The Equivalence of Inertial and Passive Gravitational Mass. *Annals of Physics*, 1964, vol. 26, 442–517.
11. Weinberg S. Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity. John Wiley & Sons, New York, 1972.
12. Brans C. and Dicke R. H. Mach's Principle and a Relativistic Theory of Gravitation. *Physical Review*, 1961, vol. 124 (3), 925–935.
13. Goenner H. F. M. Mach's Principle and Theories of Gravitation. In Barbour J. B., Pfister, H., Eds. Mach's Principle: From Newton's Bucket to Quantum Gravity. Birkhäuser, Boston, 1995, pp. 442–457.
14. Ciufolini I. and Wheeler J. A. Gravitation and Inertia. Princeton University Press, Princeton, NJ, 1995.
15. Jammer M. Concepts of Mass in Classical and Modern Physics. Dover Publications, New York, (1961) 1997.
16. Feynman R. P., Leighton R. B., Sands M. Lectures on Physics, Volume II: Mainly Electromagnetism and Matter. Addison-Wesley Publishing Company, Reading, Massachusetts, 1964.
17. Taylor E. F., Wheeler J. A. Spacetime Physics: Introduction to Special Relativity, 2nd ed. Freeman, New York, 1992, pp. 250–251.
18. Okun L. B. The Concept of Mass. *Physics Today*, 1989, v. 42 (6), 31–36.
19. Okun L. B. The Einstein Formula: $E_0 = mc^2$. "Isn't the Lord Laughing?". *Physics–Uspekhi*, 2008, v. 51 (5), 513–527. arXiv: physics.hist-ph/0808.0437.
20. Okun L. B. Energy and Mass in Relativity Theory. World Scientific, New Jersey, 2009.
21. Oas G. On the Abuse and Use of Relativistic Mass. arXiv: physics.ed-ph/0504110v2.
22. Oas G. On the Use of Relativistic Mass in Various Published Works. arXiv: physics.ed-ph/0504111.
23. Barut A. O. Electrodynamics and Classical Theory of Fields and Particles. Dover Publications, New York, 1980.
24. Charap J. M. Covariant Electrodynamics: A Concise Guide. The John Hopkins University Press, Baltimore, 2011.

A Space Charging Model for the Origin of Planets' Magnetic Fields

Tianxi Zhang

Department of Physics, Alabama A & M University, Normal, Alabama
E-mail: tianxi.zhang@aamu.edu

Both theoretical models and experimental results have indicated that a body surrounded by plasma is negatively charged to a potential around 2-3 times greater than the thermal potential of the ambient plasma. This potential difference shows that the body holds some extra electric charge. In this paper, we formulate an expression to compute the extra electric charge from the ambient plasma. It is shown that the total electric charge on a body basically depends on its size and the characteristics of the ambient plasma. When the body size is big or the ambient plasma is dense, the extra electric charge is large. Since all solar planets are imbedded within the solar wind plasma, they may also be charged due to the same physics. Analyzing the charging behavior of planets, we find that the solar planets are significantly charged. The circular electric currents or charge flows caused by planets's spinning produce magnetic fields. The magnetic fields predicted by the present space charge model basically agree with the measurements on the global magnetic fields of planets (including the Moon). Also, the polarity biases and reversals of planet magnetic fields are discussed. Therefore, a possible explanation for the origin of the magnetic fields of planets is proposed.

1 Introduction

The origin of the geomagnetic field has been puzzling physicists for hundreds of years. In 1600, Willian Gilbert believed that the Earth is permanently magnetized, like a giant magnet. Albert Einstein considered the origin of the geomagnetic field to be one of the five most important unsolved problems in physics. So far, tons of data on the geomagnetism have been accumulated [1]. In general, the geomagnetic field resembles the field generated by a dipole magnet located at the center of the Earth. The locations of the north and south geomagnetic poles are randomly varied and reverse each other at irregular periods [2-4]. The intensity of the geomagnetic field is transiently changed and in average about 0.5 G, which is slowly decayed. It is generally believed that the geomagnetic field is affected by various external events, such as the tides, aurora, solar flares, sunspots, and so on.

In order to explain the geomagnetic phenomena, various models have been proposed, which are conveniently classified into dynamo and non-dynamo models. As a non-dynamo model, the permanent magnetization of the Earth could not explain the polarity reversals of the geomagnetic field. The charge separation arising from the thermoelectric effect is, however, relatively small in comparison with the geomagnetic field [5]. In addition, some other effects were suggested - such as the gyromagnetic effect, the hall effect, the galvanomagnetic effect, the differential rotation effect, the electromagnetic induction by magnetic storms, and the Nernst-Ettinghauser effect, etc. [6-11].

Larmor [12-13] was the first to suggest that large astronomical bodies might have magnetic fields that arise from a self-exciting dynamo process. However, Cowling [14] showed that this disc dynamo was damped and cannot maintain such a field very long. Later, other dynamo theories were developed, such as magnetohydrodynamic dynamo, kinematic dynamo, turbulent hydromagnetic dynamo, and so on [15-22]. Although it is generally accepted today that the geomagnetic field arises from dynamo action in the Earth's liquid outer core, there is no viable hydrodynamic geodynamo model as described by McFadden and Merrill [23] because there are so many unclear parameters being included in the governing equations.

The study of the magnetic fields of planets offers the key to an understanding of the origin of the geomagnetic field. Until recently, information about the magnetic fields of planets came mostly from indirect measurements or from flyby missions. The measurements are generally sparse in both spatial and temporal distribution and only provide us a first-order picture of the magnetic fields of planets.

The solar planets can be conveniently classified into two types (type-I and type-II) according to their magnetic fields being local or global. A type-I planet has a weak global magnetic field, such as Venus, Mars, or Pluto (also the Moon). These planets are almost naked to the solar wind plasmas because of lack of (or very weak) magnetospheres [24]. Their atmospheres are usually not strong enough to sheath out the solar wind and partially ionized especially at the upstream. However, the type-II planets (including Mercury, Earth, Jupiter, Saturn, Uranus, and Neptune) have strong global magnetic fields. The solar wind plasmas are separated from these planets by their powerful magnetospheres except at their poles. The solar wind electric currents can still interact with the Earth through partially ionizing the neutral atoms in the atmosphere at the poles. The early measurements did

show that electric currents were observed in both the air and the Earth during aurora taking place [25-27].

Dynamo theorists suggested that the global magnetic fields of the type-II planets were excited due to their interior dynamo actions, which are critically dependent of the size and spin of the planets. The main reason Venus lacks a dynamo is because it spins slower than the Earth does. The reason Mars lacks a dynamo is because it is smaller than the Earth. However, Mercury probably has a dynamo action even though it is smaller and its spinning period is longer than Mars. A dynamo model may not easily answer why Mercury has a dynamo but Mars does not.

In this paper, a theoretical space charge model for the origin of the global magnetic fields of planets is proposed. The purpose of this paper is not to be against dynamo theories instead of to suggest another possibility. According to the space charge physics, a body floating in the space plasma will be charged. This phenomenon has been actually observed during space experiments. The electric charge on a large conducting spherical body is further derived. If the body is spinning, the electric charge will generate a circular electric current, which induces a magnetic field. It is shown that the induced magnetic field depends not only on the size and spin period of the body, but also on the characteristics of the ambient plasma. For very large bodies, such as planets, the induced magnetic fields could be as big as the measurements. Analyzing the magnetic fields and electric charging processes of the two types of planets results in a consistent explanation for the magnetic fields of all planets, including the Moon. The polarity biases and reversals of the planet magnetic fields are also discussed with this model. In addition, it should be noted that this model has not included the effects of atmosphere, body motion, and plasma instabilities on current collection. The relative motion between the body and the environmental plasma was shown to increase current collection along the magnetic field lines [28]. The field aligned current-driven instabilities was shown to greatly heat charged particles [29-30] and hence can also increase the current collected by the body.

2 Space Charging

Experimentally and theoretically, it has been shown that a satellite moving (or floating) in space plasmas itself becomes usually negatively charged, since the number of electrons incident on its surface is greater than the number of ions [31] (see Figure 1a). The absorption of electrons and ions essentially depends on the size of the body, the surface potential, the material properties of the body, and the state of the ambient plasma. In some special cases, a body may be positively charged.

The amount of electric charges and the absolute value of potential increase as long as the number of electrons and the number of ions being absorbed on its surface are not identical. Since the increasing potential slows down the electric

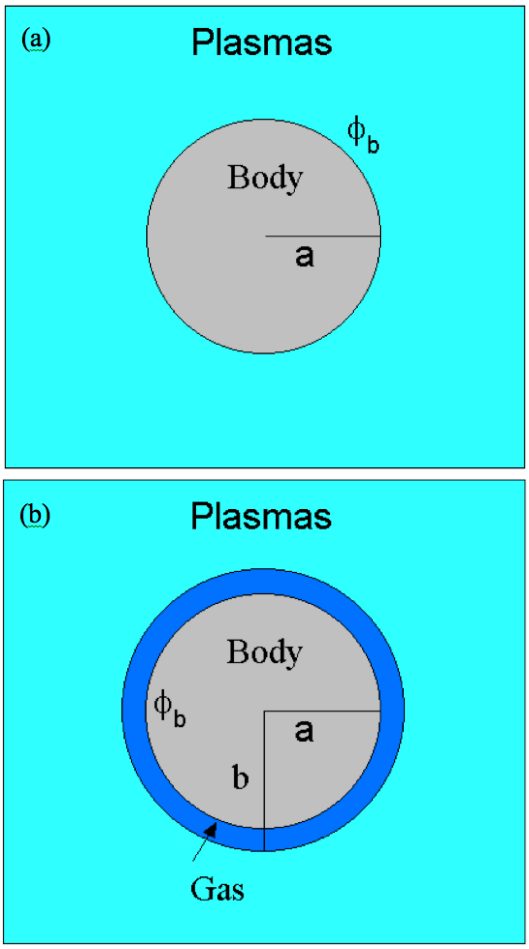

Fig. 1: Schematic diagrams for space charge model. (a) Without a neutral gas layer; (b) with a neutral gas layer.

accumulating processes, an "equilibrium" state for charge accumulating (i.e. a state in which the total electric current incident on the body is equal to zero) is finally attained if the ambient plasma is vast. In this situation, the electric potential at the surface of the body, is determined by

$$\phi_0 \simeq -\frac{k_B T_e}{e} \ln \sqrt{\frac{2k_B T_e}{\pi m_e V_0^2}} = -\alpha \phi_{th}. \qquad (1)$$

Here the potential at infinite distance (or outside the plasma sheath) has been chosen to be zero; k_B is the Boltzmann constant; e is the proton electric charge; the subscript e refers to the electron species; T_e is the electron temperature; m_e is the electron mass; V_0 is the velocity of the body relative to the ambient plasma - which is generally much larger than the ion thermal velocity and less than the electron thermal velocity; ϕ_{th} is the thermal potential which is defined by $k_B T_e/e$; and α is the factor which is given by the nature logarithm in Eq. (1). For example, we consider that a spherical body is

moving in the ionosphere plasma. The thermal potential is $\phi_{th} \sim 0.11$ V if $T_e \sim 0.1$ eV is chosen for the plasma. The factor α is $\alpha \sim 2.58$ if $V_0 = 8$ km/s is chosen for the body. Then the electric potential at the body surface is estimated as $\phi_0 \sim -0.3$ V, which is in agreement with space measurements. For a motionless or slowly moving body (i.e., $V_0 = 0$ or much smaller than the ion thermal velocity), the result is $\phi_0 \sim 2.57$ V [32].

If the body is separated from the plasma by a thin layer of neutral gas (see Figure 1b), the region of neutral gas will get extra electrons since the electrons incident into the neutral gas region are more than the ions. For a quasi-neutral plasma the number of electrons entering the neutral gas in a unit time through a unit area is determined as the electron flux ($n_0\bar{v}_e/4$), which is much greater than that of ion ($n_0\bar{v}_i/4$). Here n_0 is the number density of electron (or ion) of the quasi-neutral plasma; $\bar{v}_e$ and $\bar{v}_i$ are the mean velocity of electrons and ions, respectively. These extra electrons will diffuse toward the body because both the electron density and the electric potential have gradients. That is to say, the body will be charged. The total amount of electric charge distributed on the body and within the neutral gas should be generally greater than that without the gas layer. In present study, we limit our analyses in cases of very thin layer and hence ignore the effects of neutral gas on the charge of the body.

If the electric potential distribution is given, the electric charge on the body can be obtained. For a spherical body within a medium (including free space, dielectric medium, plasma, etc.), the density of electric charge distributed on the body surface is given by

$$\sigma_b = -\epsilon \left.\frac{d\phi}{dr}\right|_{r=a}, \tag{2}$$

where a is the radius of the body; ϕ is the electric potential distribution; r is the radial coordinate; and ϵ is the dielectric permittivity. For a static (or slowly moving) electrically conducting body, the density of electric charge on the surface is constant. Hence, the total electric charge of the body is

$$Q_b = -4\pi\epsilon a^2 \left.\frac{d\phi}{dr}\right|_{r=a}, \tag{3}$$

In the free space, a spherical body with $a = 1$ m and $\phi_0 = -0.2$ V will be charged to $Q_b = 4\pi\epsilon_0 a\phi_0 \sim -2 \times 10^{-11}$ Coulomb.

If the medium is plasma, however, the relationship between the total electric charge and the electric potential of the body will be complex. The total electric charge or the numbers of electrons and ions being absorbed by a body essentially depends not only on the potential and size of the body, but also on the state of plasma. In this case, the electric potential distribution must be generally determined through solving the Poisson equation,

$$\nabla^2\phi = -\frac{1}{\epsilon}\sum_{j=i}^{j=e} n_j q_j. \tag{4}$$

For a body with size much greater than the Debye length, however, the potential near the body can be approximately obtained only by solving the one-dimensional Poisson equation,

$$\frac{d^2\phi}{dr^2} = -\frac{n_{e0}e}{\epsilon}\left(e^{-\phi/\phi_{th}} + e^{\phi/\phi_{th}}\right). \tag{5}$$

Here the Boltzmann number density distributions have been applied for both electrons and ions. Integrating Eq. (5) one times with respect to r, we obtain

$$\frac{d\phi}{dr} = \sqrt{\frac{2n_{e0}e\phi_{th}}{\epsilon}\left(e^{-\phi/\phi_{th}} + e^{\phi/\phi_{th}} - 2\right)}. \tag{6}$$

At the surface of the body (i.e. at $r = a$), it becomes

$$\left.\frac{d\phi}{dr}\right|_{r=a} = \sqrt{\frac{2k_B T_e n_{e0}}{\epsilon}\left(e^{-\alpha} + e^{\alpha} - 2\right)}. \tag{7}$$

By substituting Eq. (7) into Eq. (3), we obtain a formula to estimate the electric charge of a large conducting body floating in space plasmas

$$Q_b = -4\pi a^2\sqrt{2\epsilon_0 k_B T_e n_{e0}\left(e^{-\alpha} + e^{\alpha} - 2\right)}, \tag{8}$$

where the dielectric permittivity has been replaced to that of free space, since we do not consider a very dense plasma. Therefore, the body is in general to be negatively charged. The amount of charge on the body depends not only on the size and potential of the body but also on the temperature and density of the ambient plasma.

Now we consider the case in which a body is moving relative to the ambient plasma. When the velocity of the body is in the range of, $v_{Te} \gg V_0 \gg v_{Ti}$, the number density of ions near the body is no longer the Boltzmann distribution, where v_{Te} and v_{Ti} are the thermal velocities of electrons and ions. In the upstream, the number density of ions is not interfered by the body if the body surface does not reflect particles. In the downstream, however, the number density of ions is almost zero since the ions slowly respond to the motion of the body. In this case, the total electric charge on the body surface (or Eq. 3) is similarly derived as

$$Q_b = -2\pi\epsilon_0 a^2\left(\left.\frac{d\phi^F}{dr}\right|_{r=a} + \left.\frac{d\phi^R}{dr}\right|_{r=a}\right), \tag{9}$$

where ϕ^F and ϕ^R are the electric potential distributions in the upstream and downstream, respectively. The electric potential distributions can be determined by

$$\frac{d^2\phi^F}{dr^2} = -\frac{n_{e0}e}{\epsilon_0}\left[1 - \exp\left(\frac{\phi^F}{\phi_{th}}\right)\right], \tag{10}$$

$$\frac{d^2\phi^R}{dr^2} = \frac{n_{e0}e}{\epsilon_0}\exp\left(\frac{\phi^R}{\phi_{th}}\right). \quad (11)$$

By integrating both Eq. (10) and Eq. (11) with respect to r once, we obtain the electric fields at the surface as

$$\left.\frac{d\phi^F}{dr}\right|_{r=a} = \sqrt{\frac{2k_BT_en_{e0}}{\epsilon_0}(e^\alpha - \alpha)}, \quad (12)$$

$$\left.\frac{d\phi^R}{dr}\right|_{r=a} = \sqrt{\frac{2k_BT_en_{e0}}{\epsilon_0}(e^\alpha - 1)}. \quad (13)$$

By substituting Eq. (12) and Eq. (13) into Eq. (9), we obtain the total electric charge of the body as

$$Q_b = -2\pi a^2\sqrt{2\epsilon_0k_BT_en_{e0}}\left(\sqrt{e^\alpha - \alpha} + \sqrt{e^\alpha - 1}\right). \quad (14)$$

This expression gives a value much greater than that from Eq. (8) if $\alpha \ll 1$. When α is not small, however, the result from Eq. (14) approaches that from Eq. (8).

If the body is spinning, the electric charge on the body surface will generate an electric circular current. This current then induces a magnetic field with poles on the spinning axis. The maximum value of the magnetic field is derived as

$$B = -\frac{\pi}{4}\mu_0\frac{Q_b}{\tau}, \quad (15)$$

where μ_0 is the permeability of free space, $\mu_0 = 4\pi \times 10^{-7}$ H/m; τ is the spin period of the body; and Q_b is given by either Eq. (8) or Eq. (14) according to the motion of the body.

Since the circular current is induced by the self-rotation of a charged body other than by the electric charges moving on the body, the magnetic field induced by the circular current (Eq. 15) is independent of the conductivity of the body surface. If the body surface is made of insulate (i.e., infinite conductivity) material, the density of electric charge will not be constant. In this case, we need to integrate Eq. 2 on the entire body surface to obtain the total charge. It is generally believed that all solar planets are not made of insulate materials. Therefore, this magnetic field formula (Eq. 15) can be generally employed to predict the induced magnetic field of a self-rotated large conducting body, such as an orbit satellite, the Moon, and the solar planets. The required parameters are the radius of the body, the spinning period of the body, the velocity of the body relative to plasma flow, the electron temperature of the ambient plasma, and the non-perturbed density of the ambient plasma. The induced magnetic field will be great when the body is large, the spin is fast, and the plasma is dense. According to the presented model the Mercury magnetic field could be greater than the Mars magnetic field because the solar wind plasma around Mercury is much denser than that around Mars.

For an orbit satellite with a conducting spherical surface, the typical required parameters are, $a = 1$ m, $T_e = 1500$ K, $n_{e0} = 10^6$ cm^3, and $V_0 = 8$ km/s. Substituting T_e and V_0 into Eq. (1) we show that the conducting satellite is charged to a potential equal to $\sim -2.6\phi_{th}$; that is, $\alpha \sim 2.6$. Substituting a, T_e, n_{e0}, and the value of α into Eq. (8) (or Eq. 14) we obtain the electric charge of the satellite, $Q_s \simeq 2 \times 10^{-8}$ Coulomb, which is much larger than that in the free space. Furthermore, if the satellite is self-rotated with a spin period, $\tau_s = 1$ seconds, the induced magnetic field, from Eq. (9), will be $B_s = 2 \times 10^{-11}$ Gausses, which is quite small. It should be noted that the rotation of the satellite does not significantly affect the ambient plasma because the linear speed at the surface due to body rotation is much smaller than the thermal velocities of ions and electrons. The presented model does not include the magnetic field effect on the body charging (or current collection) process. If the magnetic field or the body electric potential is not high, such effect is negligible [33].

3 The magnetic fields of the Moon and planets

Now, employing the space charging model proposed above, we study the magnetic fields of the Moon and planets. The predictions on the magnetic fields are compared with the measurements.

According to the space charge model, the magnetic field of a body is determined by giving the five parameters: the size and spin period of the body, the density and temperature of the plasma, and the velocity of the plasma flow. Table 1 shows the interplanetary conditions for the Moon and planets [34-35]. The solar wind velocity, density, and temperature are shown in the third to fifth column. The magnetic field of the solar wind near each planet and the distance between the Sun and each planet are also shown in this table (see the sixth and second columns). The radii and spin periods of the Moon and planets are shown in the second and third columns in Tables 2 and 3.

As mentioned in the Introduction, the Moon and type-I planets (e.g. Venus, Mars, and Pluto) are almost naked to the

Table 1: Interplanetary properties: Distance to the Sun $L_{\rm Sun}$ (AU), Solar Wind Velocity $V_{\rm SW}$ (km/s), Density $n_{\rm SW}$ (cm^{-3}), Temperature (10^4 K), and Magnetic Field (nT).

Planet	$L_{\rm Sun}$	$V_{\rm SW}$	n_{SW}	T_{SW}	B_{SW}
Mercury	0.4	430	50	20	35
Venus	0.7	430	14	17	10
Earth (or Moon)	1	430	7	15	6
Mars	1.5	430	3	13	3
Jupiter	5.2	430	1/4	9	1
Saturn	9.6	430	1/16	7	1/2
Uranus	19	430	1/50	6	1/4
Neptune	30	430	1/160	5	1/7
Saturn	39	430	1/200	4	1/10

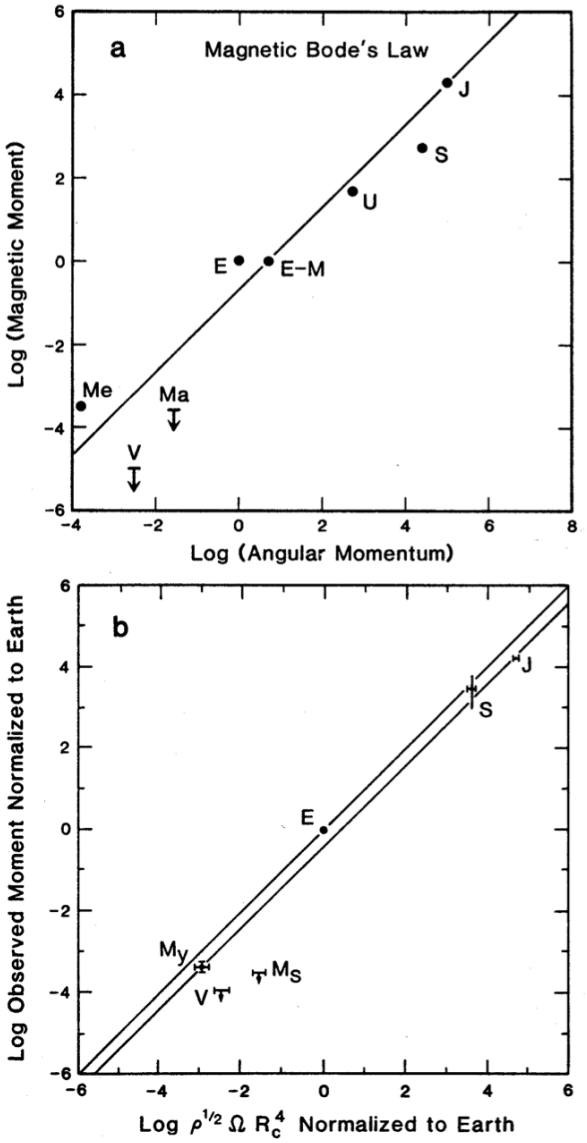

Fig. 2: (a) Relationship between planetary magnetic moment and angular momentum by the original magnetic Bode's law (Russell 1987). (b) Relationship between planetary magnetic moment and the core radius by the dynamo theory-based scaling law (Busse 1976.

solar wind plasmas (actually, they have very thin and weak gas (or atmosphere) layers). The condition, $v_{Te} \gg V_0 \gg v_{Ti}$, which is used for the deduction of Eq. 14, is generally satisfied for the Moon and the type-I planets. In this case, the ambient plasma is the solar wind, which has speed $\sim$ 400 km/s and temperature $T_e \sim T_i \sim 10^5$ K. It is not difficult to show that the solar wind speed is much smaller than the electron thermal speed but much greater than the ion thermal speed. Therefore, the space charging model proposed in the previous section (i.e. Eq. 14) can be directly employed to quantitatively predict their present magnetic fields. For the Moon, the required five parameters are, $a = 1.738\times10^6$ m, $T_e \sim 1.5\times10^5$ K, $n_{e0} \sim 7$ cm^{-3}, $V_0 = 430$ km/s, and $\tau_M = 2.36 \times 10^5$ seconds. At first, using T_e and V_0, we show, from Eq. 1, that the Moon is charged to a potential equal to $\sim -\phi_{th}$ (that is, $\alpha \sim 1$, with this value of α, the Eq. 14 predicts a result without a significant difference from Eq. 8). Then, substituting a, T_e, n_{e0}, and the value of α into Eq. 8 (or Eq. 14), we can obtain the electric charge of the Moon, $Q_M \sim -640$ Coulomb. Finally from Eq. 9, the induce lunar magnetic field (B_M) predicted by the present model is about $B_M \sim 3$ nT. The measurements actually indicate that the intensity of the global magnetic field of the Moon does not exceed 2 to 3 nT (Table 2; [36]).

For Venus, the prediction on the magnetic field by the present model is about 6 nT, which agrees with the measurements. Space experiments indicated that the intrinsic value of the magnetic field at the surface of Venus could not be greater than 5 nT [37].

For Mars, the prediction on the magnetic field by the present model is about 200 nT. In the 1970s, the soviet Mars 3 and 5 probes measured a field about 30 - 60 nT near the equator, at periapisis (at an altitude of 1500 km) [38-40]. Since the magnetic field of Mars on its surface is several times greater (for the Earth, the factor is $\sim 2-4$) than that measured at an altitude of 1500 km, the Mars' magnetic field could be as big as 150 nT, which also agrees with the present model prediction.

We also predict the magnetic field for Pluto although we have not had any measurement available so far. Based on the present model, the Pluto's magnetic field is estimated to be about 0.1 nT, which is ordinarily the same as the magnetic field of the solar wind there. Therefore, the predictions by the space charge model on the magnetic fields of the Moon and the type-I planets basically agree with the measurements (see Table 2, [36-40].

For the type-II planets (such as the Earth), however, we cannot directly obtain the present magnetic fields from Eqs. 1, 8, and 9 because the solar wind plasmas are separated from these planets by their strong magnetospheres. But, we can apply the present model to estimate the ancient magnetic fields of planets if the characteristics of the initial solar wind are known. The following gives some analyses for the type-II

Table 2: Model predictions on B for the type-I planets including Moon and Pluto in comparison with data B_0.

Planet	R (km)	τ (10^5 s)	B_0 (nT)	B (nT)
Venus	6055	210	≤ 5	6
Moon	1738	23.6	≤ 3	3
Mars	3398	0.886	~ 150	200
Pluto	1150	5.519	$\sim$	0.1

planet magnetic fields based on the evolutionary characteristics of solar system. In the next section, the type-II planet magnetic fields are further discussed through considering the polar aurora plasmas as their charging sources. If so, the present model can still be used and predicts results closer to the measurements.

It is widely believed that the Sun went through FU Orionis and T-Tauri phases of evolution [1, 41]. A T-Tauri (in the pre-main sequence) star is partially characterized by violent outbursts of material, very strong magnetic field, and an increased luminosity of about six magnitudes. Observations actually indicated very massive winds from these early-type stars [42]. Preliminary results from the studies of meteorites and lunar rocks also indicated that the average solar wind speed might have been considerably greater some $3-4\times10^9$ years ago [42,43].

Thus, it is reasonable to assume that the Sun initially emitted a strong solar wind. During that time period, all our planets were greatly charged from such massive solar wind plasmas and induced magnetic fields with different intensities due to their different sizes and rotation speeds. If the initial solar wind is $\sim 10^3 - 10^6$ times denser than the present solar wind, the ancient (or initial) magnetic fields are some tens to thousands times greater than the present fields for the Moon and the type-I planets. For the type-II planets, the ratios of the ancient fields to the present fields are in the range of $\sim 1-100$. Thus, the planets with small size and slowly spinning (such as the Moon and type-I planets) also excited considerably great intrinsic magnetic fields, which probably had magnetosphere-like structures during early periods. However, their magnetic fields are easily decayed as the solar wind becomes weak due to their weak abilities to maintain such fields. Large, fast spinning planets (such as the type-II planets) developed very strong magnetic fields and formed powerful magnetospheres - which are also decayed, but relatively more stable than the type-I planets, because they last a longer time in the decaying process.

Observations show that the planet's magnetic field is stronger if its magnetosphere is bigger. According to the presented model, the denser the ambient plasma is, the more charge the body is charged, which is proportional to the induced magnetic field. For the type-II planets (e.g. the Earth), the nearest ambient plasma is the plasmasphere (ionopshere) or the aurora plasma in the pole regions. For these plasmas (see [44]), the electron or ion density is $\sim 10^5$ to 10^6 cm^{-3} which is much denser than the solar wind plasma. The electron or ion temperature is $\sim 10^3$ to 10^4 K. In these regions, most of ions are O^+, which has a thermal velocity around 1 km/s, which is much less than the minimum speed ($\sim$ 8 km/s) for a particle to escape out by overcoming the Earth gravitation. Therefore, the Earth's (as well as other type II planets') gravity may maintain its magnetic field (or magnetosphere) through trapping the particles of plasmasphere or plasma in the aurora regions. The magnetic field itself also helps the planet to trap the particles of magnetosphere. The electrons can be trapped by the ions although the electron thermal velocity may be greater than the minimum escaping speed. Within a relative stable solar wind, the value of the magnetic field or the size of the formed magnetosphere actually depends on the planet gravity. The bigger the gravity is, the stronger the magnetic field is or the bigger the magnetosphere forms if the other parameters are the same.

The results predicted by the present model are very high in absolute values under the assumption that the ancient solar wind density varied in the range of 10^3-10^6 times denser than the present value. During such a long time interval, the planets' magnetic fields were greatly decreased when the solar wind density was greatly decreased. For the type-II planets, we have compared (in the following several paragraphs) the relative results predicted by the present model on the ancient magnetic fields of planets with the measurements and found a good agreement between them.

The fourth column of Table 3 shows the measurements of the magnetic fields for the type-II planets, which are normalized by dividing the geomagnetic field. A 300 nT magnetic field was measured for Mercury [45]; a 15 Gausses magnetic field at the north pole was measured for Jupiter [46]; and orderly $\sim$ 1 Gausses' magnetic fields were measured for Saturn, Uranus, and Neptune [47]. The fifth column of Table 3 shows the predictions of the ancient magnetic fields for the type-II planets, which are normalized by the ancient geomagnetic field. Comparing the fourth column with the fifth column of Table 3, we found that the normalized ancient magnetic fields of planets predicted by the space charging model basically agree with the present field measurements [45-48]. The Saturn's magnetic field (or magnetosphere) could be decayed more than the Jupiter's probably due to the lower gravity (or density) of Saturn.

The decays of planet magnetic fields were probably affected by their gravitation. It is reasonable to assume that a planet with large gravity has more power to maintain its magnetosphere through trapping its particles. To consider such gravity effect, we propose a formula for the present magnetic field of a type-II planet by introducing an arbitrary coefficient,

Table 3: Model predictions on B/B_e for ancient magnetic field for the type-II planets in comparison with data B_0/B_e for present magnetic field.

Planet	R (km)	τ (10^5 s)	B_0/B_e	B/B_e
Mercury	2439	51	$\sim 1/100$	1/130
Earth	6371	0.864	≤ 1	1
Jupiter	71600	0.354	~ 30	45
Saturn	60000	0.368	~ 2	13
Uranus	25600	0.621	~ 1	1
Neptune	24765	0.567	$\sim 1/2$	1/2

$f(g)$, to the Eq. 15 as

$$B = -\frac{\pi}{4}\mu_0 f(g)\frac{Q_b}{\tau}, \tag{16}$$

where g is the gravity at the planet surface and Q_b is the body charge of the planet, which is given by either Eq. 8 or Eq. 14. Then, the magnetic moment of the planet can be derived as

$$M = -\frac{2\pi}{3}a^2 f(g)\frac{Q_b}{\tau}. \tag{17}$$

It can be seen that the magnetic moment of the planet is proportional to a^4 because of $Q_b \propto a^2$. It is also proportional to the square root of the solar wind pressure and inversely proportional to the planet spin period. On the other hand, the magnetic Bode's law also called the Shuster or the Blackett hypothesis established that the magnetic moments of the planets were proportional to their angular moments (see Figure 2a and [49-50]). The scaling law predicted that the planet magnetic moments were proportional to the rotation rate times the fourth power of the core radius (see Figure 2b and [51]).

In order to compare the results predicted by the present space charging model with the predictions by either the magnetic Bode's law or by the scaling law, we plot our model predictions on the magnetic moments of the type-II planets versus the observations in Figure 3. The magnetic moments from both the model predictions and the observations are normalized to the Earth and are shown in log scales. Figure 3a has not included the gravitation effect and Figure 3b gives the results with the gravitation effect by assuming that the coefficient is linearly proportional to the gravity ($f(g) \propto g$). The observation data are from [1].

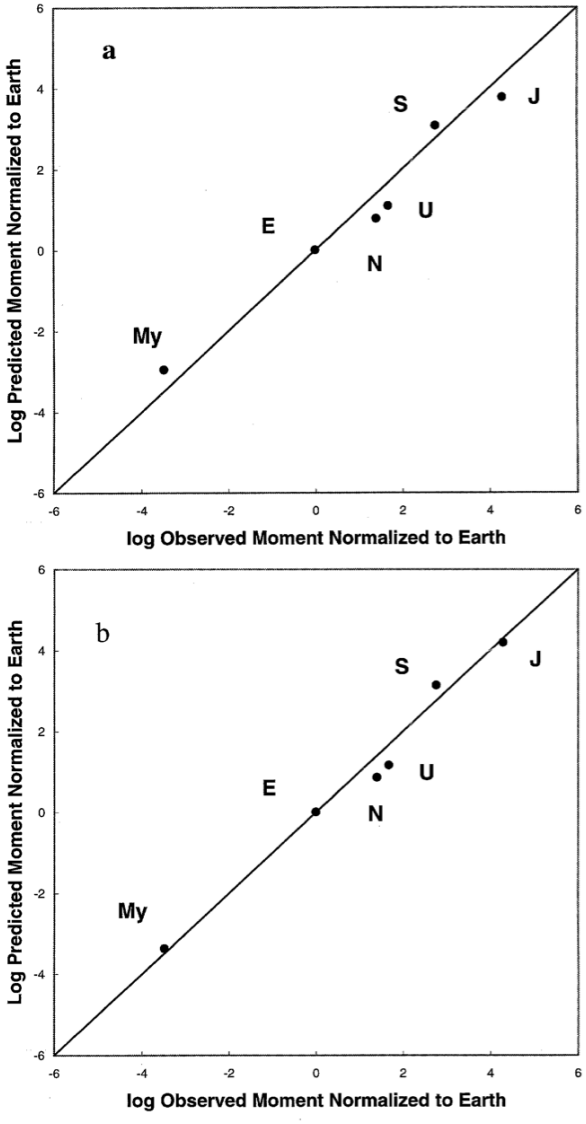

Fig. 3: Planetary magnetic moments normalized to the Earth predicted by the space charge model versus those from measurements. (a) Without the gravity effect; (b) with the gravity effect.

4 Discussions and Conclusions

In this section, we briefly discuss the following items: 1) current collection of planets with magnetospheres and 2) the polarity biases and reversals of the magnetic fields of planets. Then, we give our conclusions of this study.

Although the Earth and other type-II planets are not completely naked to the solar wind plasmas, their poles are widely opened to the outer space due to the double funnel magnetic structures. The solar and interstellar winds as well as the energetic particles can easily, through the magnetic field lines (or double funnels), come down into the polar regions of the planets to excite and to ionize the gases near the surfaces. This is the phenomena of aurora. The aurora plasmas are much denser than the solar wind plasma. The density of a typical aurora plasma could be as high as $\sim 10^5$ to 10^6 cm^{-3} which is much denser than the solar wind plasma with density less than ~ 100 cm^{-3} [44]. Therefore, these planets are probably charged at their poles especially during aurora taking place. The early experimental measurements showed that electric currents were actually observed in the air and in the Earth while the aurora was taking place [25-27]. The correlation between the Earth current and the geomagnetic activity was also found. It is interesting that if we consider the aurora plasma as the source plasma to charge the Earth, the present model predicts a result very closer to the measurement.

For the Earth, observation records show that the aurora events asymmetrically occur at the two (i.e. North and South) poles [52]. The Northern aurora events are generally more frequent and intense than the Southern aurora events. The reason is probably due to that the spinning geomagnetic field lines drift the entering (or coming down) electrons apart from the axis of spinning at the North but towards the axis of spin-

ning at the South. That is to say, the charging process at the North is faster than that at the South. This difference leads to an electric current from pole to pole. If the conductivity is different from place to place (or non-uniform) on the Earth's surface, the electric current from pole to pole will not be uniformly distributed on the Earth's surface. This polar current and the circular current will generate a total magnetic field, which biases from its rotation axis. Both the biases and the value of the induced magnetic field are transiently changed because the space charging process is transient.

The observations indicated that the geomagnetic field varies in two (long and short) time scales. In the long time (usually greater than about 100 years) scale, the field strength is decreased and the biased angle (or the orientation) of the field also changes in a certain regulation (see [1] and reference therein). On the other hand, the geomagnetic field changes transiently or in a short time scale [53]. The presented model do predict a magnetic field with such kinds of variations because the solar wind plasma transiently (short time) changes and slowly (long time) decays its plasma density. That is, according to the presented model, the planet magnetic fields should have the two time scale variation behaviors. For the type-I planets (including the Moon), the transient changes of the fields are significant because they are directly charged from solar wind plasmas (or they get extra electrons directly from the solar wind plasmas). For the type-II planets, however, the transient changes of the fields may not be significantly affected by the variation of the solar wind parameters because they do not get extra electrons directly from the solar wind plasmas. The global field does not significantly change because it is impossible for the huge magnetosphere to follow the changes of the solar wind even with the daily and season effects.

According to the present model, the original magnetosphere is arisen due to the proposed mechanism for the unmagnetized body. The unmagnetized body collected extra electric charges from the initial solar wind and formed a strong magnetosphere. If there were no solar wind later, the originally formed magnetosphere would have not existed for such a long time because of the charge being quickly released. In fact, the solar wind only slowly becomes weak. It resists (slows) the releasing of the body charge through refilling some electric charge to the body. This refilling process is actually the current collection process of a magnetized body, which can collect extra electric current (or charge) along its field lines (or at its pole regions). Therefore the energy source of the magnetosphere (or the planet magnetic field) is the solar wind. The gravity of the body also helps to maintain the magnetosphere through trapping its particles as we have discussed above.

The present model predicts that all the solar planets (including the Earth and the Moon) are negatively charged. This conclusion is in agreement with measurements if we analyze the orientation of the magnetic field and the spin direction for each planet. On the other hand, space experiments have indicated that a spacecraft could be positively charged when it has a special environment (e.g. when it goes to a great distance) [54]. Thus if the Earth becomes positively charged due to some special solar wind conditions, the orientation of its magnetic field will be reversed. But how and in what special conditions the Earth becomes positively charged is open for further study.

By the way, it should be noted that there are really a lot of current systems in the planet's magnetosphere, such as currents on the magnetospheric boundary, magnetotail currents, the ring currents, the field-aligned currents and so on. Since any plasma current will locally form a return current in the plasma, it will not have a significant contribution to the planet's global magnetic field.

In summary, we have developed a theoretical model for the origin of the magnetic fields of planets. According to the space charging physics, we have shown that a body spinning within plasma is charged and generates a dipole magnetic field. The field intensity depends on the size, the spinning speed of the body, and the state of the ambient plasma. For a large and fast spinning body in dense and hot plasma, the generated magnetic field is big. The model predictions on the present magnetic fields of the Moon, Venus, Mars and Pluto agree with the measurements; and the relative magnetic fields of Mercury, Earth, Jupiter, Saturn, Uranus, and Neptune predicted by the present model also agree with the measured data. Furthermore, this model offers an understanding of the polarity biases and reversals of the planets' magnetic fields, and hence may provide a new possible explanation for the origin of the magnetic fields of planets.

Acknowledgement

This work was supported by the NASA research and education program (NNG04GD59G), NASA EPSCoR program (NNX07AL52A), NSF CISM program, Alabama A & M University Title III program, National Natural Science Foundation of China (G40890161), and NSF REU program (PHY-1263253 & PHY-1559870).

Submitted on June 3, 2019

References

1. Merrill R.T., McElhinny M.W., McFadden P.L. The magnetic field of the Earth (Academic Press, Inc.), 1996.
2. McFadden P.L. *et al*. Reversals of the earth's magnetic field and temporal variations of the dynamo families. *Journal of Geophysics Research*, 1991, v. 96, 3923–3933.
3. Mary C., Courtillot V.A. Three-Dimensional Representation of Geomagnetic Reversal records. *Journal of Geophysics Research*, 1993, v. 98, 22461–22475.
4. Cortini M., Barton C.C. Chaos in geomagnetic reversal records: A comparison between Earth's magnetic field data and model disk dynamo data. *Journal of Geophysics Research*, 1994, v. 99, 18021–18033.
5. Merrill R.T., McElhinny, M.W., Stevenson D.J. Evidence for long-term ,asymmetries in the Earth's magnetic field and possible implictions for

dynamo theories. *Physics of the Earth and Planetary Interiors*, 1979, v. 20, 75–82.

6. Barnett S.J. Gyromagnetic effects, theory and experiments. *Physica*, 1933, v. 13, 24–58.
7. Vestine E.H. The earth's core. *Transactions of the American Geophysics Union*, 1954, v. 35, 63–72.
8. Inglis D. R. Theories of the Earth's Magnetism. *Reviews of Modern Physics*, 1955, v. 27, 212–248.
9. Chatterjee J.S. The Crust as the Possible Seat of Earth's Magnetism. *Journal of Atmosphere and Terrestrial Physics*, 1956, v. 8, 233–239.
10. Stevenson D. Planetary magnetism. *Icarus*, 1974, v. 22, 403–415.
11. Hibberd F.H. The origin of the Earth's magnetic field. *Proceedings of Royal Society of London - Mathematical and Physical Sciences*, 1979, v. A369, 31–45.
12. Larmor J. Possible Rotational Origin of Magnetic Fields of Sun and Earth. *Electrical Review*, 1919, v. 85, 412.
13. Larmor J. How Could a Rotating Body such as the Sun become a Magnet. *Reports of British Association*, 1919, v. 87, 159–160.
14. Cowling T.G. The magnetic field of sunspots. *Monthly Notices Royal Astronomical Society*, 1933, v. 94, 39–48.
15. Elsasser W.M. Induction effects in terrestrial magnetism part I. Theory. *Physical Review*, 1946, v. 69, 106–116.
16. Bullard E.C. The magnetic field within the Earth. *Proceedings of the Royal Society of London*, 1949, v. A197, 433–453.
17. Bullard E.C., Gellman D. Homogeneous Dynamos and Terrestrial Magnetism. *Philosophical Transactions of the Royal Society of London*, 1954, v. A247, 213–278.
18. Roberts G.O. Spatially periodic dynamos, *Philosophical Transactions for the Royal Society of London*, 1970, v. A266, 535–558.
19. Roberts G.O. Dynamo action of fluid motions with two-dimensional periodicity. *Philosophical Transactions for the Royal Society of London*, 1972, v. A271, 411–454.
20. Gubbins D. Theories of the geomagnetic and solar dynamics. *Reviews of Geophysics and Space Physics*, 1974, v. 12, 137–154.
21. Bullard E.C., Gubbins D. Generation of magnetic fields by fluid motions of global scale. *Geophysical and Astrophysical Fluid Dynamics*, 1977, v. 8, 43–56.
22. Moffatt H.K. Magnetic field generation in electrically conducting fluids (Cambridge University Press, New York), 1978.
23. McFadden P.L., Merrill R.T. Inhibition and geomagnetic field reversals. *Journal of Geophysics Research*, 1993, v. 98, 6189–6199.
24. Bagenal F. in Solar System Magnetic Fields (E. R. Priest, Ed., D. Reidel Publishing Company, Holland), 1986, pp. 224–254.
25. Hessler V.P., Wescott, E.M. Correlation between Earth-current and geomagnetic disturbance. *Nature*, 1959, v. 184, 627.
26. Fleming J.A. Terrestrial Magnetism and Electricity (J. A. Fleming, Ed., Dover Publications, Inc., New York), 1949, pp. 49–58
27. Helsley C.E. Magnetic properties of lunar dust and rock samples. *Science*, 1970, v. 167, 693–695.
28. Zhang T.X. *et al.* Current collection by tethered satellite TSS-1R. Seminar Presentation at Center for Space Plasma and Aeronomic Research, 1999.
29. Zhang T.X. Solar 3He-rich events and ion acceleration in two stages. *Astrophysical Journal*, 1995, v. 449, 916–929.
30. Zhang T.X. Solar 3He-rich events and electron acceleration in two stages. *Astrophysical Journal*, 1999, v. 518, 954–964.
31. Al'pert Y.L., Gurevich A.V., Pitaevskii L.P. Space physics with artificial satellites (Consultants Bureau, Now York), 1965, 63–105.
32. Herr J. L., K. S. Hwang, and S. T. Wu, *Proceedings of the 26th AIAA Plasmadynamics and Lasers Conference* (San Diego), 1995.
33. Parker L.W., Murphy B.L. Potential buildup on an electron-emitting ionospheric satellite. *Journal of Geophysics Research*, 1967, v. 72, 1631–1636.
34. Slavin J.A., Holzer R.E. Solar wind flow about the terrestrial planets. I - Modeling bow shock position and shape. *Journal of Geophysics Research*,1981, v. 86, 11401–11418.
35. Russell C.T., Baker D.N., Slavin J.A. The magnetosphere of Mercury, in Mercury (F. Vilas, C. R. Chapman, and M. S. Matthews, Eds.,The University of Arizona Press, Tucson), 1988, pp. 514–561.
36. Colburn D.S., in The Moon (S. K. Runcorn and H. C. Urey, Eds., D. Reidel Publishing Company, Dordrecht, Holland), 1971, pp. 355–371.
37. Burgess E., Venus, an Errant Twin (Columbia University Press, New York), 1985, pp. 47–48.
38. Mutch T.A. *et al.* The Geology of Mars (Princeton University Press, New Jersey), 1978, pp. 39–40.
39. Russell C.T. The magnetic field of Mars - Mars 3 evidence re-examined, *Geophysical Research Letters*, 1978, v. 5, 81–84.
40. Russell C.T. The magnetic field of Mars - Mars 5 evidence re-examined, *Geophysical Research Letters*, 1978, v. 5, 85–88.
41. Taylor S.R. Solar system evolution: A new perspective (Cambridge Univ. Press, Cambridge U.K.), 1992, pp. 58–59.
42. Holzer T. E., in the Solar system plasma physics, V.I, (C. F. Kennel, L. J. Lanzerotti, and E. N. Parker, Eds., North-Holland Publishing Company, Amsterdam-New York-Oxford), 1979, pp. 101–176.
43. Heymann D., in The solar output and its variation (O. R. White et al. Eds., Colorado Associated University Press, Boulder), 1978, pp. 405–427.
44. Al'pert Y.L. Space plasma (Cambridge University Press, Now York), 1983.
45. Connerney J. E. P. and N. F. Ness, in Mercury (F. Vilas, C. R. Chapman, and M. S. Matthews, Eds., The University of Arizona Press, Tucson), 1988, pp. 494–513.
46. Encrenaz T., Bibring T.P., Blanc M. The solar system (Spring-Verlag Berlin Heidelberg), 1990, pp. 211–214.
47. Burgess E. Uranus and Neptune (Columbia University Press, New York), 1988, pp. 82–83.
48. Bergstralh J.T., Miner, E.D. in Uranus (J. T. Bergstralh, E. D. Miner, and M. S. Matthews Eds., The University of Arizona Press, Tucson), 1991, pp. 3–25.
49. Luhmann J.G., Russell C.T., Brace L. H., Vaisberg O.L. The Intrinsic Magnetic Field and Solar-Wind Interaction of Mars, *in Mars* (The University of Arizona Press, Tucson & London), 1992, pp. 1090–1134.
50. Russel C.T. Planetary magnetism. In *Geomagnetism*, (J. A. Jacobs ed., Orlando Academic Press), 1987, v. 2, 458–523.
51. Busse F. H. Generation planetary magnetism by convection. *Physics of the Earth and Planetary,* 1976, v. 12, 350–358.
52. Laundal K.M. Ostgaad N. Asymmetric auroral intensities in the Earth's Northern and Southern hemispheres. *Nature*, 2009, v. 460, 491–493.
53. Matsushita S., Campbell, W.H. Physics of geomagnetic phenomena (Academic Press INC., New York), 1967.
54. Whipple E.C. Potentials of surfaces in space. *Reports on Progress in Physics*, 1981, v. 44, 1197–1250.
55. Chapman S. Solar plasma geomagnetism and aurora. Gordon and Breach, Science Publishers, Inc., New York, 1964.

On the Fluid Model of the Spherically Symmetric Gravitational Field

Alexander Kritov
E-mail: alex@kritov.ru

The radial flow within the frame of analogue hydrodynamic approach to gravitational field with spherical symmetry is reviewed. Such alternative models of gravity, for example the river model of black holes and the analogue gravity, do not satisfy the continuity equation for the radial fluid flow. The presented model considers a case of incompressible fluid with non-zero source-sink field that can reconcile the continuity equation with the analogue gravity. Based on modelling of a fluid parcel's evolution with time, three cases are reviewed resulting in the Schwarzschild, the Schwarzschild-de Sitter (SdS) and the Schwarzschild-Anti de Sitter(AdS) metrics. The parameters of the model are exactly determined. The model can support a view on the de Sitter cosmology and can serve as its alternative interpretation via such hydrodynamic approach.

1 Introduction

General Relativity (GR) is a widely accepted theory of gravitation. However, in spite of its mathematical beauty and concordance with experiments, as it is well known, it also has a few difficulties: first of all, it is still problematic to merge GR with quantum mechanics; secondly, GR is not fully sufficient in explaining few observable effects in the cosmology (such as rotation curves of the galaxies); and lastly it is not a singularity-free theory. In this article an alternative approach to gravitation based on the fluid/aether model is reviewed.

Such interpretations (not dismissing GR) always existed in parallel, starting from Lenz and Sommerfeld who reported his ideas in Lectures on Theoretical Physics [12] in 1944. In the 1960s, a number of authors discussed this topic following Lenz's idea, see [10, 11]. The approach uses Special Relativity (SR) only to derive the same results as GR [3–5,7,9]. Even if this model still captures the interest of the researchers, it is not widely accepted, and usually is considered through the prism of a "heuristic" approach as it was reviewed in [13].

Such four-vector model of gravity describes a spherically symmetric gravitational field via the Lorenz invariant four-potential which are the same as the components of four-vector "aether" velocity

$$v^\alpha = \left(\frac{\phi}{c}, v^r, v^\varphi, v^\theta\right) \tag{1}$$

where ϕ is the scalar gravitational potential *, and

$$v = \sqrt{\frac{2Gm}{r}} \tag{2}$$

is the radial velocity as measured by co-moving observer given for the case of a static, non-rotating mass m without charge and $v^\varphi = v^\theta = 0$. The velocity in case of the Kerr-Newman metric is obtained in [6], and in case of the de Sitter metric is reviewed in [3]. According to such approach the curvature of spacetime is the consequence of movement of some medium (or even space itself [2]). The concept implies that *something moves and therefore space curves*, [4–6]. Due to this motion the special relativistic length contraction leads to spatial curvature in gravity and the special relativistic time dilation causes time dilation in gravitational field respectively.

The Schwarzschild metric written in the $(-+++)$ sign convention generated by radial flow is given by

$$ds^2 = -c^2\left(1 - \frac{v^2}{c^2}\right)dt^2 + \left(1 - \frac{v^2}{c^2}\right)^{-1} dr^2 + r^2\, d\Omega^2 \tag{3}$$

where $d\Omega^2 = \sin^2\theta\, d\phi^2 + d\theta^2$ and the coordinate velocity is given by (2). Even if such model fully suffices to describe all effects of GR, it has two drawbacks: first, it is based on the abstract concept of moving space and does not hypothesize about the nature of what moves. It should be *something* that moves instead of nothing. Secondly, it is applicable to spherically symmetrical fields only. The second point is not as solid as the first one, because most of the objects in the universe demonstrate spherical symmetry, especially in the physics of elementary particles where the phenomena of gravitation originates.

2 The analogue gravity and its problem with the hydrodynamic continuity equation

Though, even if the ideas for a fluid theory of the gravitation were reported before [16], recently, as a continuation and generalization of such approach, the analogue gravity model was proposed [1, 14, 15]. It is based explicitly on *fluid hydrodynamics*, and it uses the acoustic metric for a moving fluid in general form (not only for spherically symmetric case) as

$$g^{\mu\nu} = \frac{\rho}{c}\begin{bmatrix} -\left(c^2 - v^2\right) & \vdots & -v^j \\ \cdots\cdots & . & \cdots\cdots \\ -v^i & \vdots & \delta^{ij} \end{bmatrix}. \tag{4}$$

*For example, the reader may check that such effective potential given by $(v_0 c)$ (its second term of the Taylor series) leads to the correction of Newtonian potential and to the same result for the anomalous perihelion precession of Mercury as GR.

In spherically symmetric case it suggests that density of the fluid should change as $r^{-3/2}$ and therefore the conformal factor appears as in the acoustic metric as

$$ds^2 \propto r^{-3/2}\left[-c^2\left(1-\frac{v^2}{c^2}\right)dt^2+ \right. \\ \left. +\left(1-\frac{v^2}{c^2}\right)^{-1}dr^2+r^2d\Omega^2\right]. \tag{5}$$

Then it creates an issue for the metric itself. The suggested workaround [1] is to represent the fluid density as perturbation $\rho=\rho_0+\rho'$ i.e. as linearized fluctuations around the background value. This is good to model the metric in approximation but again the first term does not satisfy the continuity equation.

It should be noted that such value for the velocity (2) in the frame of the fluid analogue model of gravity is not derived from any hydrodynamic equation. Moreover the inflow through the sphere of radius r as $4\pi r^2 b = r^{3/2}$ is clearly incompatible with the continuity equation. The presented approach suggests to resolve the conformal factor problem in the analogue gravity by conjecturing the fluid's *constant density* and sink-source term in the continuity equation which represents an evolution of fluid parcel's volume with time in the Lagrangian frame .

3 The continuity equation for the model

Let's consider an ideal inviscid isentropic fluid. In Lagrangian co-moving frame of reference the use of relativistic equation of the continuity is not *required* and also because, as discussed in [4], the metric in the co-moving frame is flat. In case of presence of sink-source term the equation of continuity in Lagrangian frame is

$$\frac{\partial\rho}{\partial t}+\rho\nabla\cdot(\vec{v})=\sigma \tag{6}$$

where σ is the sink-source term. In case of constant density ρ_0 it reduces to

$$\nabla\cdot(\vec{v})=\frac{\sigma}{\rho_0}=\frac{\partial\dot{V}}{\partial V} \tag{7}$$

where the rate of volume production per time within a control volume was denoted as $\dot{V}$. Let's now consider the spherically symmetric case and take some volume with radius r. Using the Gauss-Ostrogradsky theorem then

$$4\pi r^2 v(r)=\dot{V}(r)=\frac{1}{\rho_0}\int_0^r\sigma(r)4\pi r^2 dr \tag{8}$$

where $\dot{V}$ represents the total volume integral of sink-sources σ within a sphere of radius r. So the radial velocity can be obtained from (8) as

$$v(r)=\frac{\dot{V}}{4\pi r^2}. \tag{9}$$

In (9) the rate of volume production is a function of time in Lagrangian frame $\dot{V}(t)$, or in Eulerian frame is a function of only radial distance $\dot{V}(r)$ respectively, and the flow is stationary.

It is important to make note on a sign of the velocity (2). The approach is valid for both – for positive and negative values of the velocity (2) because it comes to the metric (3) as squared value. Many authors treat the river model of gravity with radial flow going in inward direction to the center of gravity. However, in the present model it is considered opposite – the outward flow of the fluid and the positive sign for velocity (placing coordinate center at the point mass) which means that the flow is decelerated going from the point mass center and has also negative acceleration.

4 The linear model, the Schwarzschild metric

Let's now consider the point mass m and the spherical coordinate center is placed in m. The point mass m emits the volume parcels V_n of the fluid at some constant rate ω_m with initial position $r=0$ and time $t=0$. The parameter ω_m is denoted in such way because of an assumption that it depends on the property of point mass itself or even may be linearly proportional to the value of point mass m. So every time interval

$$\Delta t = 1/\omega_m, \tag{10}$$

one n^{th} parcel of the fluid V_n appears near the point m and no initial velocity is considered. Following the above, let's assume that every parcel V_n further grows linearly with time in its respective Lagrangian frame as *

$$V_n=\omega V_0 t \tag{11}$$

where $V_0=m_0\rho_0$ and ω are some external constants which do not depend on the property of point mass, and ω is in the same way linearly proportional to a parameter m_0. Then the total number of produced parcels during time t is

$$n=\omega_m t. \tag{12}$$

So, the volume of n^{th} parcel in row is given by

$$V_n=\frac{\omega}{\omega_m}V_0 n. \tag{13}$$

Importantly, time in Lagrangian frame (local co-moving frame of every fluid's parcel) is synchronized with time of the observer resting at infinity (see [5] for more details on this). So, the time interval given by (10) is the same in the co-moving frame of parcel as well as in the reference frame of point mass.

In order to find $\dot{V}$ within a sphere of some fixed radius r, first a total volume produced by sum of all such parcels has

*For simplicity one can imagine the emitted volume parcels V_n as growing spherical bubbles, though fluid parcels have no actual form.

to be defined. Summation of (13) yields

$$V(t) = \sum_{1}^{n} V_n = \frac{\omega}{\omega_m} V_0 \frac{n^2}{2} = \frac{1}{2}\,\omega_m \omega V_0 t^2 \qquad (14)$$

where an approximation that $n \approx n+1$ for a relatively big number of parcels was used. Taking time derivative and substituting into (9) leads to

$$v = \frac{dr}{dt} = \frac{\omega_m \omega V_0 t}{4\pi r^2}\,. \qquad (15)$$

Solving this differential equation for $r(t)$ one can find the equation of motion for the fluid as

$$r(t) = \left(\frac{3\omega_m \omega V_0 t^2}{8\pi} + c_1\right)^{1/3} \qquad (16)$$

where c_1 is an arbitrary constant and represents initial position of parcel at time $t = 0$ which has to be zero, so c_1 is zeroed. Expressing $t(r)$ from (16) and substituting this into the original equation (15) results in the fluid velocity $v(r)$ in Lagrangian frame as

$$v = \frac{dr}{dt} = \left(\frac{1}{6\pi}\frac{\omega_m \omega V_0}{r}\right)^{1/2}\,. \qquad (17)$$

So as a result, the radial velocity is inversely proportional to the square root of the radial distance as (2), which reproduces the Schwarzschild metric. But still, the unknown parameters in the expression are to be determined.

The fluid acceleration is

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + (v\nabla)\,v\,. \qquad (18)$$

For a stationary radial flow the acceleration is given only by the convective term, therefore

$$a = \nabla\left(\frac{v^2}{2}\right) = -\frac{1}{12\pi}\frac{\omega_m \omega V_0}{r^2}\,. \qquad (19)$$

This acceleration is negative for the positive value of the velocity (17), and as the coordinate center was placed in the center of mass m, it means that the flow is decelerated in outward direction. However, as it was noted above, the corresponding metric (3) remains the same regardless of the velocity sign.

5 The volume conversion relation and the uncertainty principle

Let's introduce the volume V_m such as

$$V_m = \frac{m}{\rho_0} \qquad (20)$$

where m is the mass of the point source. And let's assume that ω_m represents de Broglie wave frequency of the mass m, and m_0 is given by the uncertainty principle with rigorous factor of two (where it originates because of the non-commutativity of the quantum operators [8]) as

$$m_0 c^2 = \rho_0 V_0 = \frac{1}{2}\,\hbar\omega\,. \qquad (21)$$

This means that the fluid parcel's mass m_0 is not observable during the time ω^{-1}. Then

$$V_m \omega = 2\omega_m V_0\,. \qquad (22)$$

Further this expression will be referred as the volume conversion relation with the exact factor of two. Therefore (17) becomes

$$v = \left(\frac{\omega^2}{12\pi\rho_0}\frac{m}{r}\right)^{1/2}\,. \qquad (23)$$

Regarding the mass-energy conservation, the point mass m does not act as actual source studied in classical fluid dynamics, because at time $t = 0$ an outgoing parcel has zero volume $V_n = 0$ and zero mass accordingly, therefore there is no actual mass flow from the point mass m. The linear mass growth of a parcel is also governed by the uncertainty principle and it is not observable during the time ω^{-1}.

6 The hyperbolic model, the SdS metric

Presumably the linear dependency of $V_n(t)$ in the model above can be just an approximation of some unknown odd function and the linear function of t in (11) represents just a first term of its Taylor series. Choosing to test the hyperbolic sine one may assume that V_n changes with time in its respective Lagrangian frame as

$$V_n = V_0 \sinh(\omega t)\,. \qquad (24)$$

Considering that time in co-moving frame of parcel now is not synchronized with time running at the clock of the observer at rest at infinity, but the time coordinate transform is given by

$$t' = \frac{1}{\omega}\sinh(\omega t) \qquad (25)$$

where t' is proper time in co-moving parcel's frame.

Following the same procedure, as in the previous model, the total number of produced fluid parcels during time t is given by (12). And the volume of n^{th} parcel in row is given by

$$V_n = V_0 \sinh\left(\frac{\omega}{\omega_m} n\right)\,. \qquad (26)$$

The sum of all such parcels provides the total volume produced by time t as

$$V(t) = \sum_{1}^{n} V_n = V_0 \frac{\sinh^2\left(\frac{n}{2}\frac{\omega}{\omega_m}\right)}{\sinh\left(\frac{1}{2}\frac{\omega}{\omega_m}\right)} \qquad (27)$$

where $n \approx n+1$ for a relatively big number of parcels. The value of $\sinh\left(\frac{1}{2}\frac{\omega}{\omega_m}\right)$ is very small and can be easily approximated without a loss of precision as $\frac{1}{2}\frac{\omega}{\omega_m}$ *. Then, using trigonometric identity and t instead of n let's rewrite (27) in simpler form as

$$V(t) = \frac{\omega_m V_0}{\omega}\left(\cosh(\omega t) + 1\right) \tag{28}$$

where factor $1/2$ disappears because of the trigonometric conversion. Taking time derivative and using the volume conversion relation (22) it becomes

$$\dot{V} = \frac{1}{2}\,\omega V_m \sinh(\omega t)\,. \tag{29}$$

With the use of (9) the differential equation is

$$v = \frac{dr(t)}{dt} = \frac{\omega V_m \sinh(\omega t)}{8\pi r(t)^2}\,. \tag{30}$$

Solution for $r(t)$ provides the equation of motion as

$$r(t) = \left({r_0}^3 + \frac{3V_m \cosh(\omega t)}{8\pi}\right)^{1/3}\,. \tag{31}$$

Applying boundary condition as $r = 0$ when $t = 0$ the equation of motion becomes simply

$$r(t) = \left(\frac{3V_m}{8\pi}\right)^{1/3}\left(\cosh(\omega t) - 1\right)^{1/3}\,. \tag{32}$$

Expressing the hyperbolic sine from this and then substituting it into (30) leads to

$$v(r) = \left(\frac{V_m\omega^2}{12\pi r} + \frac{\omega^2 r^2}{9}\right)^{1/2} \tag{33}$$

or with use of the definition of V_m (20) the resulting radial velocity is

$$v(r) = \left(\frac{\omega^2}{12\pi\rho_0}\frac{m}{r} + \frac{\omega^2 r^2}{9}\right)^{1/2}\,. \tag{34}$$

So the hyperbolic model leads to the same radial velocity as in the previous model (23), but with the additional term. Using (18) the fluid acceleration is

$$a = -\frac{\omega^2}{24\pi\rho_0}\frac{m}{r^2} + \frac{\omega^2 r}{9}\,. \tag{35}$$

7 Determination of the model parameters

The association of the first term in (35) with Newtonian gravitational acceleration allows expressing the value for fluid density via ω as

$$\rho_0 = \frac{\omega^2}{24\pi G}\,. \tag{36}$$

*For example for the proton mass such approximation would give an error of order less than 10^{-40}.

Then substituting ω from this into the second term of (35) gives the repulsive acceleration as

$$a_{rep} = \frac{8\pi}{3}\rho_0 G r\,. \tag{37}$$

This term can be also treated as the Newtonian gravitational force from uniformly distributed mass that has the equation of state $p = -\rho c^2$ and satisfies stress-energy equivalent

$$\rho_0 + \sum_i \frac{p_i}{c^2} = -2\rho_0 \tag{38}$$

as given in [13, see the expressions (45–46)]. Assuming the constant density ρ_0 (36) is equal to the critical density, the value for ω can be defined via the Hubble constant as

$$\omega = 3H\,. \tag{39}$$

And the repulsive acceleration as given by (35) is

$$a_{rep} = H^2 r = \frac{c^2\Lambda}{3}r\,. \tag{40}$$

The radial velocity of the fluid (34) based on (3) and using (39) leads to

$$\begin{aligned} ds^2 = &-\left(1 - \frac{2Gm}{c^2 r} - \frac{H^2 r^2}{c^2}\right)c^2 dt^2 + \\ &+ \left(1 - \frac{2Gm}{c^2 r} - \frac{H^2 r^2}{c^2}\right)^{-1} dr^2 + r^2 d\Omega^2 \end{aligned} \tag{41}$$

that corresponds to the Schwarzschild-de Sitter metric for the hyperbolic model.

8 The harmonic model, the Schwarzschild-AdS metric

Using the sine function in (25) which could be treated as a simple harmonic oscillation of a fluid parcel volume $V_n(t)$. Following the same procedure (substituting sinh() with sin() instead) it is easy to see that the result would be the same as it was in previous model (34) but with a difference in sign of the second term

$$v(r) = \left(\frac{\omega^2}{12\pi\rho_0}\frac{m}{r} - \frac{\omega^2 r^2}{9}\right)^{1/2} \tag{42}$$

which with the use of (39) and (3) obviously leads to the Schwarzschild-Anti de Sitter metric.

9 Conclusions

The model results in full accordance with known metrics with exact accuracy by the coefficients based on assumptions of the volume conversion equation (22) and of the equality of the fluid density to the critical density value. The forces, the Newtonian gravitational and the repulsive cosmological, both

appear natively in the hyperbolic model. Therefore the model may support a view on applicability of the static de Sitter metric for cosmology. In presented approach the de Sitter Universe is also empty in the sense that the mass of the matter is attributed to the medium with constant density ρ_0. While the matter objects may reside statically at the fixed coordinates of the metrics (41), the space-time curvature (resulting in both attractive gravitation and repulsion) originates in a motion of the medium. The equation of state and the stress-energy of such fluid were suggested (38). However, one should be cautious to apply GR for further analysis of the solutions, because only Special Relativity is considered in the frame of the present approach.

The fluid parcels can be treated as virtual particles emitted by an elementary particle with the constant rate given by the de Broglie frequency, and on the other hand they can be considered as "growing bubbles of space". An individual parcel is not observable during the cosmological time, and its mass and volume are constrained by the uncertainty principle as shown.

The evolution of parcel's volume with time was modelled by odd functions. The odd functions have property of being asymmetric under time-reversal transformation. The requirement for such time asymmetry to generate velocities applicable to describe different metrics for gravitational field could be a topic for future study. Further analysis is required on finite boundary conditions (when a fluid parcel originates at time $t = 0$ at finite radius) and on corresponding event horizons. The temporal coordinate transform (25) as a base of the hyperbolic model, a possible correspondence of the cosmological scale factor to the proposed volume increase require further analysis.

Received on May 17, 2019

References

1. Barcelo C., Liberati S., Visser M. Analogue Gravity. arXiv: gr-qc/0505065.
2. Braeck S., Gron O. A river model of space. arXiv: gr-qc/1204.0419.
3. Cuzinatto Z. Z., Pimentel B. M., Pompeia P. J. Schwarzschild and de Sitter solution from the argument by Lenz and Sommerfeld. *American Journal of Physics*, 79:662, 2011. arXiv: gr-qc/1009.3249.
4. Czerniawski J. What is wrong with Schwarzschild's coordinates. arXiv: gr-qc/0201037.
5. Czerniawski J. The possibility of a simple derivation of the Schwarzschild metric. arXiv: gr-qc/0611104.
6. Hamilton A. J. S., Lisle J. P. The river model of black holes. *American Journal of Physics*, 2008, v. 76, 519–532. arXiv: gr-qc/0411060.
7. Kassner K. A physics-first approach to the Schwarzschild metric. *Advanced Studies in Theoretical Physics*, 2017, v. 11 (4), 179–212. arXiv: gr-qc/1602.08309.
8. Robertson H. P. The Uncertainty Principle. *Physical Review*, 1929, v. 32, 163–164.
9. Rowlands P. A simple approach to the experimental consequences of general relativity. *Physics Education*, 1999, v. 32 (1), 49.
10. Sacks W. M, Ball J. A. Simple Derivation of the Schwarzschild Metric. *American Journal of Physics*, 1968, v. 36, 240.
11. Shiff L. I. On Experimental Tests of the General Theory of Relativity. *American Journal of Physics*, 1960, v. 28, 340.
12. Sommerfeld A. Electrodynamics. Lectures on Theoretical Physics, Vol. III, Academic Press, New York, 1952.
13. Visser M. Heuristic Approach to the Schwarzschild Geometry. arXiv: gr-qc/0309072.
14. Visser M. Acoustic Propagation in Fluids: An Unexpected Example of Lorezian Geometry. arXiv: gr-qc/9311028.
15. Weinfurtner S. Emergent Spacetimes. arXiv: gr-qc/0711.4416v1.
16. Winterberg F. Vector Theory of Gravity with Substratum. *Zeitschrift für Naturforschung A*, 1988, v. 43 (4), 369–384.

GR = QM: Revealing the Common Origin for Gravitation and Quantum Mechanics via a Feedback Signal Approach to Fundamental Particle Behavior

Franklin Potter

8642 Marvale Drive, Huntington Beach, CA 92646 USA. E-mail: frank11hb@yahoo.com

By allowing the fundamental particles of the Standard Model to communicate via "feedback signals" within a vacuum lattice of mathematical nodes at the Planck scale, one learns that this approach toward understanding fundamental physics reveals the surprising common origin of quantum mechanics and of general relativity. This "feedback signal" approach is shown to be equivalent to the path integral approach but also the underlying reason for its success.

1 Introduction

The $GR = QM$ in the title refers to a recent suggestion [1] that perhaps the long-standing theoretical conflict between general relativity and quantum mechanics is not insurmountable. In fact, the conjecture has been that they may actually be closely related, or at least they could have the same fundamental origin.

Herein I establish the common fundamental origin for gravitation and quantum mechanics. A non-traditional approach to fundamental particle behavior is required, one that agrees with the successful effective Standard Model (SM) of leptons and quarks [2] but treats these particles as harmonic oscillators emitting and receiving scalar waves at their Compton frequencies [3]. A fundamental particle, such as an electron, communicates with the surrounding discrete vacuum lattice of mathematical nodes via these scalar "feedback signals". Therefore, a particle itself actively determines its subsequent behavior even in the absence of the SM local gauge fields.

The surprising result is that the common origin of quantum mechanics and of general relativity arises directly by simply analyzing particle behavior in sufficient geometrical detail.

2 A brief particle physics review

In this section I offer a brief review of some of the physics consequences if one considers both the internal symmetry space for defining the particle states of the SM and our (3+1)-D spacetime to be discrete spaces. Such possibilities may be necessary in order to justify (1) treating the internal symmetry space and spacetime as C^2 unitary space lattices of mathematical nodes and (2) proposing the leptons, hadrons, and electroweak (EW) bosons to be 3-D particles behaving as harmonic oscillators. If one chooses to accept these concepts outright, one can skip forward to Section 3 for the details of the feedback signal approach.

Recall that the SM describes the known local gauge interactions, color and electroweak, via its $SU(3)_C$ x $SU(2)_L$ x $U(1)_Y$ lagrangian, so I will ignore these gauge interactions in the discussion ahead. The leptons, the hadrons formed from quarks and gluons, and the EW interaction bosons $W^{\pm}$, Z^0, and γ, are the fundamental particles defined [2] in the internal symmetry space whose behavior in spacetime will be explained in terms of the feedback signal approach. That is, I am treating these three categories of fundamental particles as 3-D objects and not as point particles. The justification is provided below.

The proposed feedback signal approach can only be self-consistent if each fundamental fermion, i.e., lepton or quark, "gathers in" the immediate surrounding lattice nodes in its own unique way. That is, I assume that (3+1)-D spacetime is a discrete lattice of mathematical nodes, and a particle's collection of lattice nodes, perhaps at the Planck scale, must have a different discrete rotational symmetry for each different fundamental fermion family. These assumptions are in contrast to the same SU(2) point particle continuous symmetries for each family in the traditional interpretation of the SM.

Specifically, one finds that only discrete symmetry binary subgroups of the unit quaternion group Q, which is equivalent to SU(2), suffice, with each binary subgroup of Q having two EW isospin $\pm\frac{1}{2}$ states in each fermion family. Therefore, being binary subgroups of Q, and of SU(2) x U(1), all the mathematical machinery of the SM remains valid. Moreover, the important left-handed fermion state preference for the weak interaction is dictated by the mathematical properties of the quaternion multiplications for the weak interaction.

I have identified 3 discrete symmetry binary subgroups of Q that define the 3 physical lepton families [4–6]. They are these specific 3 binary subgroups acting in the R^3 subspace of C^2: the [332] binary subgroup for the electron family; the [432] binary subgroup for the muon family; and the [532] binary subgroup for the tau family. They are known also as the binary tetrahedral group 2T, the binary octahedral group 2O, and the binary icosahedral group 2I, respectively, and correspond to special discrete binary rotations of 3-D objects called regular polyhedrons in the 3-D real space R^3. No more lepton families are predicted because there are no more binary subgroups of Q that require a 3-D space.

The fact that Nature agrees with the 3 lepton families representing these 3 binary subgroups of Q is verified by the

first principles derivation [6, 7] of the neutrino PMNS mixing angles from their three quaternion generators by collectively mimicking the SU(2) generators, i.e., the three Pauli generators. The empirical values of the lepton mixing angles now agree within 1σ to each of these theoretical absolute values: $\theta_{12} = 34.281°$, $\theta_{23} = 42.859°$, $\theta_{13} = 8.578°$. Conceptually, this EW flavor state mixing to produce the mass states occurs because a valid renormalizable conformal field theory requires a continuous symmetry such as in the lagrangian of the SM. This lepton family mixing therefore guarantees that the 3 discrete symmetry binary subgroups defining the lepton families collectively behave as the SU(2) of the SM.

I have identified also 4 related discrete symmetry binary subgroups [4, 5, 8] that define four 4-D quark families in R^4: [333], [433], [343], and [533], corresponding to the only regular polytopes in R^4. The mathematical and physical consequences of these discrete symmetry groups for 4 quark families are discussed in Appendix A. The 4-D quarks and 4-D gluons combine according to QCD to form the 3-D hadrons, the baryons and mesons, or one can use intersection theory to establish the same results.

Note that the 3-D lepton states in R^3 and the 4-D quark states in R^4 both fit into the proposed 2-D unitary space C^2. Our (3+1)-D spacetime for discussing the particle bahavior also fits into C^2. I am assuming that the two spaces, the internal symmetry space for particle definition and spacetime for the physics behavior join together seamlessly. Therefore, this $C^2 = R^4$ space is proposed to be the one I need to consider to be discrete and composed of mathematical nodes. The nodes are equally spaced on average at the Planck scale when no fundamental particles are in existence.

Each fermion family with its own unique discrete symmetry binary subgroup has two Q or SU(2) orthogonal $\pm\frac{1}{2}$ states, but they will be mass-energy degenerate unless they form the two new physical orthogonal states of different energies as dictated by QM. Therefore, each lepton and each quark family has two weak isospin flavor states that have different mass values with a characteristic oscillation occurring between the two original mathematical states at the Compton frequency and Compton wavelength

$$\omega_C = \frac{mc^2}{\hbar}, \quad \lambda_C = \frac{h}{mc}. \tag{1}$$

For the electron, its Compton values are $\omega_C \approx 7.8 \text{ x } 10^{20}$ Hz and $\lambda_C \approx 2.4 \text{ x } 10^{-12}$ meters. Therefore, the Compton wavelength of each fundamental particle will be many orders of magnitude larger than the Planck distance of about 10^{-35} meters. Consequently, the proposed vacuum lattice structure of nodes appears to be a continuous space for the fundamental particles.

Although the effective SM lagrangian has the continuous symmetry local gauge group $SU(3)_C \text{ x } SU(2)_L \text{ x } U(1)_Y$, addons called horizontal discrete symmetry groups are now becoming acceptable alternatives for defining the lepton family states, particularly with the advent of neutrino mixing and non-zero neutrino mass states [2]. However, the discrete symmetry binary subgroups of the unit quaternion group Q that I have proposed for the leptons and quarks retain the successful predictions of the SM without the need to introduce any additional horizontal discrete symmetries to its lagrangian.

That is, all the successes of the SM have been retained by my specific discrete symmetry approach for the fermions while the geometrical sources of some of its physical properties have been elucidated. I cannot overemphasize this retention of the SM mathematical and physical properties, with perhaps the SM being a useful approximation even down to the Planck scale.

The above brief review of my discrete symmetry approach to the SM has been included in order to introduce some of the mathematical connections that propose some unconfirmed physics possibilities and also to justify using a discrete spacetime of mathematical nodes as both the origin of the fundamental fermions of the SM and as an active participant in their physical behavior. I will show how this approach leads directly to the special theory of relativity (STR), path integrals, quantum mechanics (QM), and the general theory of relativity (GTR), as explained in the discussion ahead.

3 The feedback signal approach

Spacetime itself at the Planck scale of about 10^{-35} meters could be a discrete space described by a uniform lattice of mathematical nodes. Therefore, I assume that our physical (3+1)-D spacetime agrees with a uniform lattice in the unitary space C^2 (or equivalently R^4) at or near the Planck scale and that each fundamental lepton family forms its particle states by "gathering in" lattice nodes to form its own unique discrete symmetry 3-D objects. This "gathering in" process distorts the lattice locally with the amount of lattice distortion extending outward in a decreasing manner with increasing distance, i.e., as inverse distance.

If I assume that the undistorted, uniformly spaced lattice has no net energy density, then the positive mass-energy of a fundamental particle is related to the amount of lattice distortion in some yet-to-be-determined way. I expect this mass-energy to be balanced by an equal negative energy value that retains the overall net zero energy total even for the distorted lattice. Perhaps the increased "stretch distance" between the nodes outside the particle definition volume provides negative energy that is the balancing factor for an assumed zero total energy for the Universe.

Recall that Clifford algebra and Bott periodicity [9] dictate a conjugate $R^4 = C^2$ space. In this conjugate space for anti-particles, the same mathematical properties of the uniformly spaced lattice would apply, again producing a positive mass-energy for the anti-particle states.

Each fundamental particle oscillating at its characteris-

tic frequency, its Compton frequency ω_C, is proposed to be emitting scalar waves, call them "feedback signals", into the surrounding vacuum lattice to eventually reach everywhere. The particle source could be undergoing "breathing mode" oscillations and emitting spherical waves isotropically into its environment. One must not identify these oscillations with electromagnetic waves because they are just propagating lattice distortions that allow lattice nodes to communicate with their nearest neighbors.

According to the special theory of relativity (STR), there exists a limiting speed for mass-energy transfer. I will take this maximum speed to be c, the speed of light in a vacuum, although there could be a higher speed limit if some day a photon is determined to possess a very tiny mass value.

Let a particle oscillate at its Compton frequency

$$\omega_C = \frac{mc^2}{\hbar}, \tag{2}$$

with m the particle's mass value, c the speed of light in a vacuum, and $\hbar$ being Planck's constant divided by 2π.

The feedback signals obey the standard scalar wave equation, a hyperbolic partial differential equation in three spatial variables x, y, z, and one time variable t. Its scalar function u(x,y,z,t) obeys

$$\nabla^2 u - \frac{\partial^2 u}{c^2 \partial t^2} = 0. \tag{3}$$

Solutions of this equation for spherical symmetry have no angular dependence, so the feedback signal amplitude u(r,t) depends only upon the radial distance according to

$$\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{\partial^2}{c^2 \partial t^2} \right) u(r,t) = 0. \tag{4}$$

The solutions for a single frequency ω have the form

$$u(r,t) = \frac{A}{r} e^{i(\omega t \pm kr)} \tag{5}$$

where the wavenumber k = ω/c and the peak intensity I(r) = $|A|^2/r^2$, i.e., the inverse square dependence.

This feedback signal approach requires the fundamental particle to behave as a microscopic 'antenna' moving within and communicating with the lattice and with other particles via its feedback signals. For example, the electron oscillating at ω_C = 7.77 x 10^{20} Hz disturbs the surrounding lattice at the same frequency ω_C, and this oscillatory disturbance propagates radially outward in all directions at speed c. By treating the particle as an antenna, the particle not only emits its feedback signals but also can absorb its own feedback signals returning from scatterings in the lattice environment.

I can describe the electron's oscillation in more detail. Although I have its oscillations only at the Compton frequency ω_C, such ideal behavior cannot be maintained once signals return from the environment, even when the electron is at rest. There will exist a small spread in frequency values about its Compton frequency according to Fourier analysis. Therefore, a Q value can be assigned to represent the small spread in frequency values, just as for any other harmonic oscillator. The signal emissions have a small spread in frequencies also, but for simplicity I will ignore this property unless needed for clarification purposes. Therefore, I will continue to use a single characteristic Compton frequency ω_C even though we understand that the oscillator does not have an infinite Q value.

The lattice nodes act as a *transponder* to the feedback signals, absorbing and immediately emitting them equally in all directions for all frequencies, all amplitudes, and with no phase shift. That is, each small volume element in the lattice must absorb some of the incident feedback signals and then emit immediately the feedback signals at the original frequency into all directions isotropically. One can think of a single lattice node or of a specific collection of lattice nodes acting together as a transponder, but considering the same type of transponder everywhere for simplicity.

If one wishes to introduce a non-zero phase shift at each transponder, then a simple modification could be to have the phase shift value be the same for all the transponders and be independent of the feedback signal frequency. Either constraint can be eliminated for a more complicated vacuum lattice. I have chosen the simplest assumption of no phase shift and equal response for all frequencies and amplitudes.

I had initially allowed the feedback signals to have an arbitrary velocity v_0. However, I learned that if one lets the speed of the feedback signals $v_0 = c$, the speed of light in a vacuum, then this simple feedback signal approach permits the direct derivation of the phenomena and equations of special relativity, general relativity, and quantum mechanics, with all of them agreeing with the present theories. The biggest surprise occurred when I learned that general relativity and quantum mechanics would then have the same fundamental origin.

In the sections ahead I will use many parts of my original 1982 attempt toward establishing this feedback signal approach as a viable approach but with some added updates here and there to provide a 21st century perspective. The identification of the gravitational interaction is one recent addition.

4 Single particle behavior at uniform velocity

Let a lone fundamental particle, such as a single electron in the Universe, be a 3-D physical harmonic oscillator oscillating at its Compton frequency ω_C with its antenna-like behavior emitting its feedback signal oscillations into the surrounding discrete lattice of uniformly spaced mathematical nodes, perhaps separated by the Planck distance of about 10^{-35} meters. As far as the electron is concerned, with its Compton wavelength of about 2.4 x 10^{-12} meters, the lattice appears to be continuous. Likewise for all other particles composed of leptons and of quarks, i.e., the hadrons, as well as the interaction bosons of the SM.

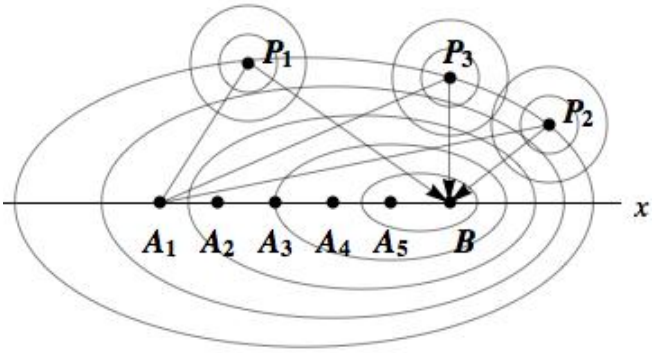

Fig. 1: Feedback signals are emitted by an electron at its previous successive equal-phase positions A_i. P_1 and P_2 are two of numerous transponders in the surrounding 3-D space on equal-phase ellipsoids for the signals from electron position A_1 only. This uniform velocity electron has moved forward at $0.866c$ to B where the returning feedback signals define its present location.

Either way, having a discrete space or a continuous space, the oscillations of the particle will appear as feedback signals traveling in the surrounding space R^3 (the subspace of R^4 and C^2) and progress through the space at the constant velocity c with decreasing amplitude as the radial distance from the particle increases. Why require a decreasing amplitude? Because we must consider the concept of energy conservation associated with these outgoing and incoming feedback signals.

For simplicity only, I ignore at first the "permanent" space distortion of the lattice caused by the formation and presence of each fundamental particle. Therefore, the feedback signals propagate through a lattice in which the average lattice node spacings remain the same separation distance everywhere. Later on I will remove this restriction in order to discuss gravitational effects between two fundamental particles.

Both a coordinate space and a momentum space description of this feedback signal approach is considered. Single particle behavior in coordinate space is shown in Fig. 1. If the electron had been at rest, then all the positions A_i and position B would coincide and the ellipses would be circles centered at B to exhibit the spherical symmetry. However, this electron has been moving at a uniform velocity in the +x direction and is now at location B receiving feedback signals from the transponders P_i everywhere in space. The surrounding ellipsoids are equal-phase locations for the outgoing feedback signals emitted by the electron at previous equally-spaced successive electron positions A_i for $i = 1,2,3,4,5$.

In this lab frame as the electron moves by, the diagram shows three feedback signal rays, from A_1 to P_1 to B, from A_1 to P_2 to B, and from A_1 to P_3 to B, of equal total length that have feedback signals arriving at B exactly in-phase with the particle oscillation when the particle arrives at B. These rays are a few examples of the feedback signals that have been emitted isotropically into 4π solid angle by the particle when at A_1.

Only a specific subset of all the equal-phase ellipsoids are shown in Fig. 1. Note also that each larger ellipsoid represents a lesser signal amplitude at the transponders along the ellipsoid, being a further distance away from the source, and that all feedback signals returning from the same 3-D ellipsoid have identical amplitudes and phases because their total path distances are equal. Because the transponders in space are everywhere, all emitted signals will eventually reach one of them. I will later explain how all the multiple scattering paths from the A_i to B are related to the path integral concept considered by R.P. Feynman in his approach to quantum mechanics and classical mechanics [10].

If the particle has just come into existence, then the signals will have not reached very far into the surrounding space. In almost all practical cases the particle has existed for a time long enough so that the signals will have permeated to tremendous distances and an approximate steady-state condition will have been established, with the outgoing and incoming signal amplitude totals approximately matching at the particle's new location B.

Recall that I have chosen no phase delay for the transponders. Incoming feedback signals to the transponder from any direction are immediately emitted into all directions. Their spherically symmetrical emission pattern, shown at each P_i, assumes that all space locations, and therefore all transponders, are identical, behave identically, and will "scatter" feedback signals. This ideal transponder behavior is the simplest possible for determining the subsequent behavior of the particle.

5 Frequency shifted feedback signals

The feedback signals sent forward and backward along the electron's velocity (momentum) vector in the x-direction experience frequency shifts. Signals sent in the forward direction with frequency ω_C return from those transponders at a higher frequency $\omega_C + \Delta\omega$ because the moving particle encounters the equally-spaced equal-phase maximum signal amplitudes at shorter time intervals than when the particle is at rest. That is, these returning signals at frequency $\omega_C + \Delta\omega$ are blue-shifted according to the relativistic Doppler expression

$$\omega' = \omega_C + \Delta\omega = \sqrt{\frac{1 + v/c}{1 - v/c}}\,\omega_C. \qquad (6)$$

And those feedback signals returning from transponders in the backward direction are red-shifted to the lower frequency by taking the opposite sign of the electron's velocity v.

One important consequence of this feedback signal approach is that a steady-state equilibrium can be maintained for the electron moving at a constant velocity. There is symmetric behavior in the two coordinate directions perpendicu-

lar to the velocity direction but a constant asymmetric reach in the x-direction of motion. For example, in Fig. 1 consider the outermost ellipsoid scattering the feedback signals emitted from position A_1. The backward sampling distance for a particular ellipsoid is shorter than the forward sampling distance in the environment.

In the steady-state condition for a single electron in the universe, the returning signals from all directions should not change the electron's constant velocity because there is no amplitude change in any of the returning signals, and their phases from all directions agree at the new electron position B. If there were no frequency shifts in the x-component of the feedback signal frequencies, then one might calculate the contributions by either of two methods: (1) adding up the returning signals from the rear and from the front by considering cones of equal solid angles on opposite sides of B and using elliptic functions of the second kind, or (2) adding up the returning signals along a line through B at any angle θ with respect to the velocity vector direction. Using the second method, one would add contributions along a line at angle θ to achieve

$$-\sqrt{\frac{1 + v\cos\theta_f/c}{1 - v\cos\theta_f/c}}\,\nu + \sqrt{\frac{1 - v\cos\theta_b/c}{1 + v\cos\theta_b/c}}\,\nu = 0, \qquad (7)$$

where the first term represents signals returning from the forward direction at angle θ_f and the second term returning signals from the back at angle θ_b. Because one can constrain $0 \leq |\theta| \leq \pi/2$ for the forward direction, then along the same line θ_b = -($\pi/2$ +θ_f) and the sum is always zero because the cosines have opposite signs in diagonally opposite quadrants.

However, that method does not apply for this situation. Why not? Because we must account for the frequency shifts by integrating over the surface area of each ellipsoid separately for the feedback signals returning from the forward direction and those returning from the backward direction in order to determine the net effect. In Fig. 1, one recognizes that the plane passing through points P_3 and B perpendicular to the x-axis separates the two surface parts for each ellipsoidal surface integral, thereby separating the backward returning feedback signals from the forward returning ones.

In terms of the semi-major axis b and the semi-minor axis a, the ellipsoid's eccentricity

$$\epsilon = \sqrt{(b^2 - a^2)/b^2}. \qquad (8)$$

The solid angle of the ellipsoidal cap on the right of B subtended from A_1 is

$$\Omega_{cap} = 2\pi\,(1 - \cos\theta) \qquad (9)$$

where θ is the angle between the ray from A_1 to P_3 and the x-axis. The solid angle subtended by the left side is

$$\Omega_{left} = 4\pi - \Omega_{cap} = 2\pi\,(1 + \cos\theta). \qquad (10)$$

Substituting the pertinent geometrical values, one obtains

$$\Omega_{cap} = 2\pi\left(1 - \frac{2\epsilon^3}{\sqrt{1 + 4\epsilon^6}}\right). \qquad (11)$$

These geometrical factors are multiplied by the frequencies returning from each point on the ellipsoidal surfaces. Along the x-axis one obtains:

$$\Omega_{cap}\,\omega' = 2\pi\left(1 - \frac{2\epsilon^3}{\sqrt{1 + 4\epsilon^6}}\right)\sqrt{\frac{1 + v/c}{1 - v/c}}\,\omega_C, \qquad (12)$$

and

$$\Omega_{left}\,\omega' = 2\pi\left(1 + \frac{2\epsilon^3}{\sqrt{1 + 4\epsilon^6}}\right)\sqrt{\frac{1 - v/c}{1 + v/c}}\,\omega_C. \qquad (13)$$

Substituting $\epsilon = \beta = v/c$, assuming $v << c$, and expanding the expressions in a Taylor series, their difference becomes

$$\text{Diff} \approx -4\pi\,\omega_C\,\beta\,(\beta^2 - 1) \approx 4\pi\,\omega_C\,\beta, \qquad (14)$$

i.e., proportional to the velocity v as expected, verifying that the uniform velocity will be maintained along the x-axis.

If one desires to check the result for relativistic velocities, the complete integration over the cap and the surface area remainder would be necessary. The frequency shifts can be large enough to put the returning feedback signals outside the high Q absorption curves. However, the integration verifies that the uniform velocity is maintained.

6 Inertia and Mach's principles

The idea of inertia considered in the early 1600s by Galileo and others proposed that a body maintains its state of uniform motion unless acted upon by an outside net force.

In the previous section, my feedback signal approach reveals the origin for this Law of Inertia. That is, the vacuum lattice itself plays an active and important role in maintaining the state of a particle's uniform motion. The feedback signals scatter from the transponders to arrive back in-phase to determine the particle's new location.

Information about the environment is brought back to determine the continuous behavior of the particle. Long-lived particles can establish a steady-state communication with the environment, but short-lived particles learn only transient information about their immediate environment. Fast particles near the speed of the feedback signals sample only an extremely small distance perpendicular to the trajectory direction.

The distant parts of the Universe play their role in determining the particle motion locally because feedback signals from way out there are added to the closer contributions to determine its new location. Mach's principle connecting local behavior to the influences from far reaches of the Universe therefore fits well in this feedback signal approach. The origin of the inertia concept is established.

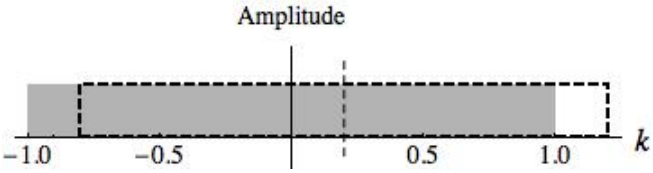

Fig. 2: The momentum-space x-component amplitude contribution at B of the returning feedback signals for the electron at rest [solid rectangle] versus the contributions of the feedback signals [dashed rectangle] in the x-direction for the electron at a uniform velocity.

7 The momentum-space description

What does the particle do with its own returning signals? And with other particle's signals, which may be at the same frequency or at other frequencies? The response to any feedback signals by the particle depends upon whether the feedback signals lie within the response range of frequencies for its inherent harmonic oscillator, meaning that feedback signals are absorbed if they lie within the absorption response curve defined by its Q value. That is, a particle is not a transponder and will be frequency selective. And, in contrast to the transponders, which maintain their initial properties forever, the future behavior of the particle can be affected.

The x-component momentum-space behavior of the electron's feedback signals is shown in Fig. 2. The gray rectangle represents the equal x-momentum contributions from all 4π solid angle for the electron at rest in the lab frame, being symmetrical about $k_x = 0$. Left to right, from $-k_x$ to 0 to $+k_x$, one has the momentum-space total amplitude contributions from the x-components of the returning feedback signals. The dashed rectangle represents the same electron moving at a constant velocity v, so this dashed rectangle is the original rectangle displaced by the x-momentum of the particle. Out-of-phase returning feedback signals will change the distribution.

8 Time asymmetry

In addition to continuous Lie symmetries, discrete symmetries are important in particle physics. Experiments in the 1950s and 1960s established both parity P and charge-parity CP violation for the weak interaction. Theoretically, one expects CPT invariance, which includes the time reversal operation T, and to this date all evidence points toward CPT conservation [2]. CP violation occurs for the weak interaction, so then T violation must occur for the weak interaction also in order to maintain CPT invariance. The mathematical source [6] of the weak interaction CP violation is simply the mathematics of products of unit quaternions in the group Q, the leptons, quarks, and weak bosons all being represented by quaternions.

This feedback signal approach to particle behavior possesses a fundamental time asymmetry, the expected T violation. Consider a free particle with its Compton frequency ω_C in uniform motion in the lab frame. To the moving particle, as we demonstrated earlier, its returning feedback signals from the forward direction are blue-shifted to a higher frequency and those returning from the backward direction are red-shifted to a lower frequency.

Now introduce time reversal via the operator T, i.e., have the electron move backwards at the same uniform velocity as if running a video backwards. The particle will be emitting bluish feedback signals in the new backward direction and their returning signals from the transponders would be red-shifted back to the original Compton frequency ω_C. The new forward emitted reddish signals will return as blue-shifted back to the original ω_C also. Therefore, the environment appears symmetrical in the forward and backward directions, so the particle should not be moving. There is a conflict with the hypothesis of time reversal symmetry. Therefore, time reversal symmetry is violated. Time reversal cannot occur in Nature.

Hence, a definite time direction is an inherent feature of the feedback signal approach. The moving particle "knows" its forward direction in the time coordinate. All particles would possess this time asymmetry property. For the anti-particles, which exist in the mathematically conjugate space to our normal space, they would also have one time direction only, forward for them but perhaps in the backward direction mathematically for us.

Consequently, time travel backwards in time would be impossible in our Universe of particles unless, perhaps, one changes all the material particles to their anti-particles that are conjectured to have the opposite time direction in the conjugate space. And time travel forward in time faster than normal would be impossible also because there would exist a conflict with the particle behavior we have established via the feedback signal approach.

9 Origin of Special Relativity

Does this feedback model of particle behavior, as developed so far, lead to the special theory of relativity (STR)? If one examines the successive series of ellipsoids shown in Fig. 1, these ellipsoids belong to a set of curves with eccentricity $\epsilon = \beta = v/c$, the ratio of the electron's velocity divided by the speed of light. Therefore, as $\beta = v/c \to 1$, then also $\epsilon \to 1$.

In order to derive the expected STR equations, two assumptions about the feedback signals must be accepted:

1. the speed of the feedback signals in all reference frames is the same constant c, and
2. the perpendicular distances are invariant.

In the laboratory frame the feedback signals from each A_i to an ellipsoidal shell and back to the electron now at B will arrive in-phase at B, the definition of the new location of the free electron. A specific path within an ellipsoidal shell

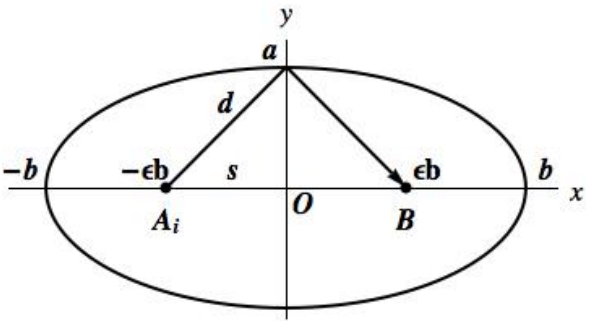

Fig. 3: Constant phase ellipsoid parameters for deriving special relativity relations in a vacuum with an eccentricity $\epsilon = \beta = v/c$.

is shown in Fig. 3. The feedback signal goes from A_i at one focus of the ellipsoid to B at the other focus in the same time that the electron goes from A_i to B via a straight trajectory through the origin O.

I can now do the standard derivation, with the feedback signals instead of with light rays. Let the lab frame be the primed frame. The perpendicular distance from O to a, the semi-minor axis distance, and back is

$$2\Delta y = 2ct, \tag{15}$$

and, using the geometrical properties of the ellipsoid,

$$2\Delta y' = 2s\left[\frac{c^2}{v^2} - 1\right]^{1/2}, \tag{16}$$

with $s = vt'$. Because the perpendicular distances in the two reference frames are equal, $\Delta y' = \Delta y$, the time intervals are related by

$$t' = \frac{t}{\sqrt{1 - v^2/c^2}} \tag{17}$$

and the distance intervals along the velocity vector in the x-diection are related by

$$\ell' = \ell\sqrt{1 - v^2/c^2}. \tag{18}$$

These relations are the fundamental equations of STR for the coordinate and time measurements. In Subsection 9.2 the relativistic energy and momentum expressions are derived. But first some geometrical properties of ellipsoids must be introduced.

9.1 Ellipsoidal geometry

In terms of the semi-major axis length b and the semi-minor axis length a, the ellipsoid's eccentricity is given by Eq. 8. If the perpendicular semi-minor axis length a is held fixed in both perpendicular directions to the x-direction as $\beta = \epsilon \to 1$, the semi-major axis value

$$b = \frac{a}{\sqrt{1 - \epsilon^2}} \to \infty. \tag{19}$$

At the same time the surface area of the ellipsoid as a prolate spheroid becomes

$$S.A. = 2\pi a^2 + 2\pi\frac{ab\,\sin^{-1}\epsilon}{\epsilon} \sim 2\pi a^2 + 2\pi ab \to \infty, \tag{20}$$

while the ellipsoid volume increases as

$$\text{Volume} = \frac{4}{3}\pi \text{ba}^2 \to \infty. \tag{21}$$

With the ellipsoids stretching out along the x-axis, the velocity direction, as a consequence of $\beta = \epsilon \to 1$, the number of in-phase ellipsoids that can "scatter" feedback signals from the A_i to B is rapidly decreasing. Or so it seems that way! As a check, consider the feedback signal that goes rearward from A_i to -b and then is scattered forward to B. If the electron's velocity $v \sim c$, then immediately after the feedback signal's emission directed toward -b comes the return feedback signal to arrive at B simultaneously and in-phase with the electron. Consequently, only a very small distance into the environment behind and sideward will be sampled to determine the electron's behavior.

The minimum sampling distance in the direction perpendicular to the x-axis might seem to be the semi-minor axis distance

$$a = \frac{ct'}{2}\sqrt{1 - \beta^2} \to 0. \tag{22}$$

However, the particle's Compton wavelength, or actually half the Compton wavelength, is the minimum sampling distance when $v \sim c$.

9.2 Energy and momentum

Using Fig. 3 again, one can determine several other important consequences in STR via the feedback signal approach. Relativistic energy and momentum can be related to the volume of the ellipsoid. If this statement is true, then the electron at rest has its mass-energy $E = mc^2$ determined by its "spherical volume" density when $\epsilon = 0$. Note that this fundamental particle volume will maintain a discrete rotational symmetry corresponding to the binary subgroup properties of each fundamental particle. So the "spherical volume" is an idealized spherical approximation in which the particle exists.

The ellipsoid volume when $\beta \ll 1$ is expressed as

$$V = \frac{4}{3}\pi ba^2 = \frac{4}{3}\pi\frac{a^3}{\sqrt{1 - \epsilon^2}} \simeq \frac{4}{3}\pi a^3\left(1 + \frac{1}{2}\beta^2 + \ldots\right) \tag{23}$$

or, when multiplied by c^2, is

$$Vc^2 = \left(\frac{4}{3}\pi a^3\right)c^2 + \frac{1}{2}\left(\frac{4}{3}\pi a^3\right)v^2 + \ldots, \tag{24}$$

which can be compared favorably to the familiar STR expansion of $m = m_0/\sqrt{1 - v^2/c^2}$ as

$$mc^2 = m_0c^2 + \frac{1}{2}m_0v^2 + \ldots \tag{25}$$

in which the second term on the right in Eqs. 24 & 25 expresses the increase of the mass-energy due to the particle's velocity, also known as the kinetic energy, and defines p = mv.

The simplest conclusion is that mass-energy is directly associated with the distorted volume of the space lattice occupied by the electron and depends upon the mass density

$$\rho(m_0) = \frac{6}{\pi}\,\frac{m_0^4 c^3}{h^3}, \tag{26}$$

which reminds us that each type of fundamental particle distorts the lattice space in its own way to pack in its unique amount of mass-energy.

But there is more to behold! The vacuum, i.e., the lattice of mathematical nodes, must contribute the energy per unit volume which can be assimilated into the moving particle to increase its total energy according to STR. Until now I have assumed that the uniformly-spaced lattice does not have energy per unit volume, which is probably correct, but now we learn that the *distorted* lattice created by the particle at rest (and when in motion) is the energy source. At this point in my earlier research in the 1970s and 1980s I realized that each fundamental particle in Nature should have a different symmetry in order to agree with my discovery of the mass-energy relation to the volume enclosed.

In 1984, by accident, I found the significant clue to the lepton family symmetries that indicated that they could be representing specific discrete symmetry binary subgroups of SU(2), i.e., the unit quaternion group Q. That is, the 3 lepton families could be representing the specific 3-D discrete symmetry binary subgroups of Q named [3,3,2], [4,3,2], and [5,3,2], and also exhibit properties and behavior that suggests that the SM is a good theory all the way down to the Planck scale with its possible discrete lattice of mathematical nodes.

10 Origin of Feynman path integrals

Physicist R.P. Feynman is credited with providing a relativistic path integral approach to quantum mechanics (QM) in the 1940s and applying this method to better understand the foundations of physics. Today, practically all areas of physics continue to use path integrals to investigate the behavior of Nature at all levels [11].

The fundamental idea behind the path integral calculation is that a particle, such as an electron, "sniffs out" all possible paths between its initial location A and its final location B. Each possible path contributes its QM amplitude and phase angle to the path integral. Most paths contribute very little to this limit of the sum because their path lengths from A to B are so long that not only are their QM amplitude values reduced significantly but also their phase values differ enough to cancel each other. Two path examples are shown in Fig. 4 that will have significantly different contributions to the amplitudes at B.

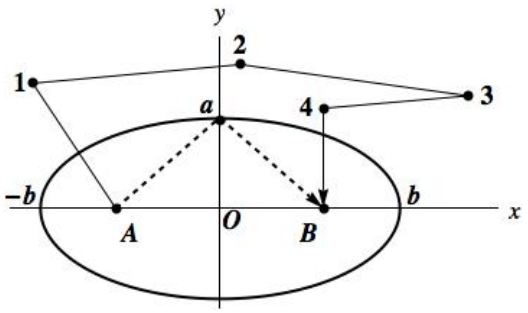

Fig. 4: Two vastly different paths from A to B: (1) Path A,1,2,3,4,B, and (2) Path A, a, B. Feedback signals travel both paths. Or, in the path integral approach, the electron "sniffs out" both paths.

The actual classical path taken will be among the paths that collectively make the biggest contribution to the path integral, because this classical path will be the one for which the nearby paths have almost the exact same contribution to the path integral. Note that this path integral approach is based upon the mathematical principle of least time, which dictates that the actual classical path will be the one for which many nearby paths have the least time difference for going from A to B. Fundamentally, the method agrees with the least action principle.

The path integral approach is a proven method that works for all of physics, quantum and classical, meaning that the path integral results agree with all the known fundamental laws of Nature. Therefore, if the feedback signal approach is the source of the path integral method, then one can explain why path integrals successfully describe all of physics! Or vice-versa!

Feedback signals are emitted by a fundamental particle into all directions and undergo multiple transponder scatterings between the initial position A of the electron and its next position B, such as the simple 5-component path in Fig. 4. All the possible paths taken by these feedback signals going from A to B can be considered collectively identical to the "sniffing" out all possible paths from A to B in the path integral approach. Each feedback signal path is then a contributor to the path integral with its specific amplitude and phase angle.

Therefore, the underlying mathematical reason why the path integral approach works so impeccably well is that fundamental particles are using feedback signals to sample their environment in order to determine their subsequent behavior. Thus, one could use path integrals as the preferred mathematical method to describe all the results of the particle feedback signal behavior.

There exist many mathematical ways to represent the path integral method. One interesting visual way [12] to represent this limit of the sum over all paths is to use equal length arrows for each path and point them in the correct phase direction in an Argand diagram shown in Fig. 5. That is, each path

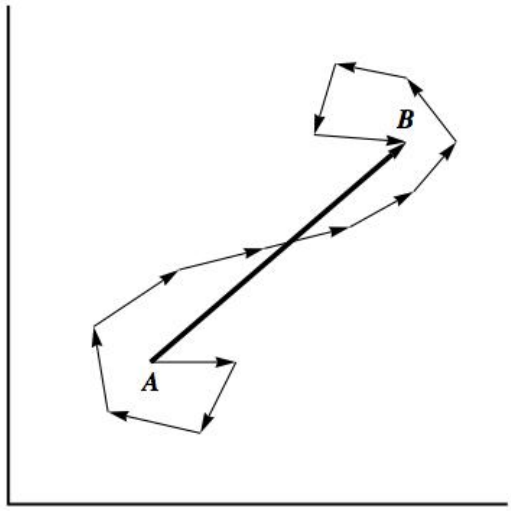

Fig. 5: Argand diagram of the phases for the different paths. Only 13 different paths are shown here, but the general idea of finding their total contribution to the amplitude is represented by the length of the arrow from A to B.

to the current position will have a different phase, therefore a different angle with respect to the horizontal real axis and the vertical imaginary axis in this complex 2-D space.

Nearby paths will have almost the same phase angle, will point in nearly the same direction, and will add a significant distance to the total vector sum in the diagram. Those arrows with opposite directions may cancel out completely. Each phase arrow is produced by a different path from the start to the current position B. The path integral amplitude is the length of the long straight arrow from beginning to end, A to B in the diagram, and the probability to be at the current position is the absolute value of its square.

In summary, each arrow also represents a feedback signal path and its phase contribution at location B, the current position of the electron. Again, one must add up all the feedback signal amplitudes arriving at B to find their total amplitude, which will depend upon the distance traveled and the phase at arrival at B. The electron position will be at the new maximum amplitude value. Therefore, we have conceptual and mathematical agreement with the path integral.

11 Origin of quantum mechanics

The rules of quantum mechanics (QM) can be derived from the path integral approach. But the path integral approach has its origin in the feedback signal approach as described above. At this point I could simply consider using path integrals to derive the 3 rules of QM. But deriving QM by the feedback signal approach provides a better "feeling" for how any particle behaves in the single slit and double slit experiments. There is no surprise because the feedback signal approach has been shown to be equivalent to the path integral approach.

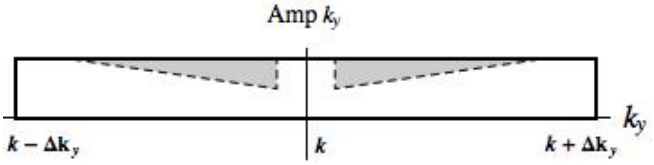

Fig. 6: While passing through the single slit the particle will experience diffraction spreading in the y direction because the feedback signals returning from the wall will produce phase shifts in the shaded regions (approximate idealized representation).

Here are those 3 rules of QM from which all its consequences can be derived [13]. But first I must recall the definition of an event in relativistic QM. A QM event is defined as a set of initial and final conditions, e.g., an electron leaves the source, arrives at the detector, and nothing else happens. The first principles of QM [i.e., the 3 rules] are:

1. Each event in an ideal expereiment is described by a complex number ψ that is called the probability amplitude, the event probability P being the square of the absolute value $|\psi|^2$.
2. When an event can occur in several alternative ways, the total probability amplitude Ψ for the event is the sum of the probability amplitudes for each way considered separately. There is an interference term $2\psi_1\psi_2$:
$$\Psi = \psi_1 + \psi_2$$
$$P = |\psi_1 + \psi_2|^2 .$$
3. If an experiment is performed that is capable of determining whether one or the other alternative is actually taken, the probability of the event is the sum of the probabilities for each alternative. The interference is lost:
$$P = P_1 + P_2.$$

Note that one does not need to actually do the measurement for this sum of probabilities to apply. Simply having the capability to do the measurement is enough to eliminate the interference terms.

11.1 Diffraction

Consider a fundamental particle moving along the x-axis approaching a narrow vertical slit extending upward along the z-axis in a solid material wall that extends to infinite distances perpendicular to the x-axis. The slit is symmetrical about the x-axis in both perpendicular directions. The particle approaches the slit from the left, goes through the slit, and recedes away from the slit to the right. One can put a "screen" of particle detectors behind the slit to measure the particle's arrival pattern.

As the particle approaches the slit the returning feedback signals define its new positions as before. Those signals re-

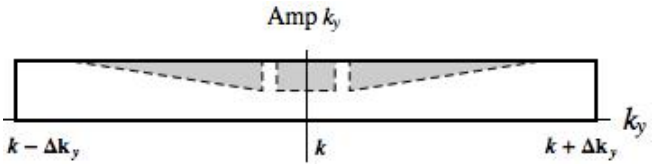

Fig. 7: After passing through the double slits the electron will experience diffraction and interference spreading in the y direction because the feedback signals returning from the wall will produce phase shifts in the shaded regions (approximate idealized representation).

turning directly from the open slit portion of the wall introduce no phase shift, and the returning signals from the volume of "empty" space on either side of the wall, front and back, also do not introduce a phase shift.

We now need to determine the phase shift effects of the wall on the behavior of the particle. Its matter content introduces phase shifts δ^a on the approach and δ^r on the recession, with the same phase shift values for each of the infinite series of feedback signal ellipsoids. The resulting amplitude values at the particle's position will depend also upon the total round trip distance.

Our concern is what happens in k-space on the back side of the slit both along the x-direction of the electron's travel and what happens in both perpendicular y and z directions. The angular distribution of the feedback signals from each ellipsoid will produce changes in the y-amplitude A(k_y) according to the actual distribution of matter around the slit. The new k_y amplitudes are shown in Fig. 6 for inside and immediately behind the slit. Within the free particle rectangular box are shaded regions for possible examples of the phase-shifted signals returning from the particles in the wall around the slit.

With left-right symmetry in the slit region itself in the y direction, there exist symmetrical amplitude decreases as shown in Fig. 6 but no net acceleration. Instead, the change in the distribution of the amplitude in k_y space leads to a symmetrical spreading of the particle according to Fourier analysis. If the wall effectively stretches to infinity, then the major contribution comes from the slit region around $k_y = 0$. One has a broadened diffraction pattern produced which has the amplitude

$$U'(y) = U(y) + 2\Delta k\, A'(k_0) \exp\left[i(\omega(k_{0y})t - k_{0y}y)\right] \times \frac{\sin\, \Delta k_y(y - v_{0y}t)}{\Delta k_y(y - v_{0y}t)}. \quad (27)$$

The term U(y) is the standard distribution in coordinate space for a free particle. The important result is the increased spread in the y-direction to produce the expected diffraction pattern, as represented by the 2nd term.

11.2 Interference

This feedback signal approach also reproduces the double slit interference pattern for the feedback signals because of the k_y momentum distribution shown in Fig. 7. In coordinate space the behavior of the feedback signals at each slit is wave-like but now one cannot determine in principle whether the particle goes through either slit because the feedback signals pass through both slits simultaneously. The amplitudes are added to produce interference before calculating the total probability.

Only when the experimental setup is such that one could determine the slit used by the particle do we get the addition of the probabilities. The mathematics tells us that whether one "looks" or not is irrelevant, but as long as one "could look", then the interference terms are absent in the probability expression.

I have explained how the particle's feedback signal behavior at a slit exactly dictates the behavior of a particle as described by QM, both for diffraction and interference. Hence, the 3 rules outlined at the beginning of this section for the first principles of QM follow directly from the diffraction and interference of the feedback signals, thereby revealing the origin of QM.

12 Origin of gravitation

Now consider the behavior of two different particles with different mass-energy values. The case of two identical particles exchanging feedback signals is discussed in Appendix B, where the connection between particle spin and quantum statistics agrees with Fermi-Dirac and Bose-Einstein behavior.

The analysis developed here first outlines the feedback signal source of the gravitational interaction. Then I discuss its agreement with the standard geometrical curvature approach to the general theory of relativity (GTR).

As an example, let's bring a muon into the environment of our electron with both particles at rest initially. I ignore their electromagnetic charge interaction, which is understood to be a local interaction described by the Standard Model, requiring the exchange of virtual photons.

Therefore, the muon has its Compton frequency about 207 times higher and a wavelength about 207 times shorter than for the electron. Thus, in Fig. 8, I cannot do justice to both particles at the same time by drawing their feedback signal ellipsoids to relative scale. Consequently, I only show different wavelength signals emitted by each, but they are not to scale.

Both particles emit their characteristic frequency feedback signals into the vacuum lattice. Each high Q particle has a nearly zero ability to absorb the signals from the other particle. Therefore, the biggest contribution to the amplitude and phase changes of the returning feedback signals comes from the lattice distortion surrounding each particle.

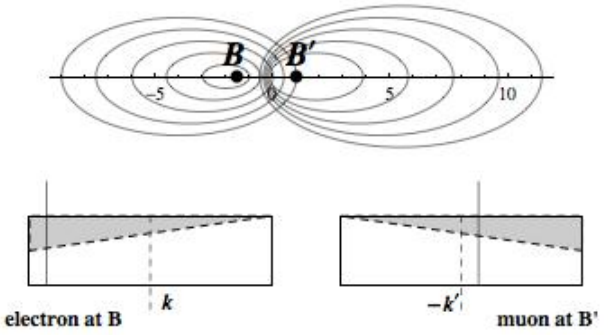

Fig. 8: As two "assumed neutral" particles approach each other, the transponders in the vacuum lattice handle both sets of feedback signals simultaneously. The signals returning from these transponders are a different phase than for the free particle. The instantaneous effects in k-space are shown in momentum space with the new k values at the dashed lines (*approximate idealized representation exagerated*).

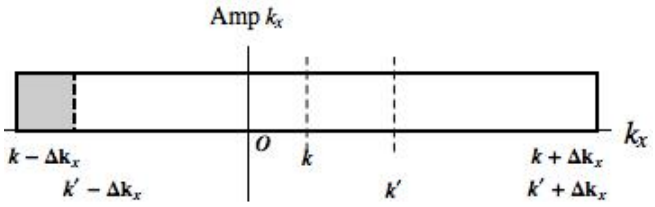

Fig. 9: Whenever a "chunk" of k-space is absent (the gray area) near k-Δ k_x, there will be an acceleration in the +x direction. Usually the feedback signals returning from the forward direction are out-of-phase, the source being the transponders around other particles in the environment ahead.

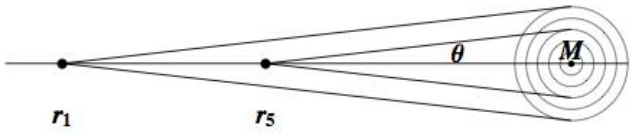

Fig. 10: As a neutral particle of mass m approaches from the left toward another neutral particle of mass M, the transponders in the vacuum lattice handle both sets of feedback signals simultaneously. Shown are paths from positions r_1 and r_5 subtending the same angle θ to the distorted space around M as "seen" from the approaching particle. The feedback signals returning from these two rings of transponders around M return with a different phase than for the free particle without the presence of M. (approximate representation exagerated).

Meanwhile, the transponders in space continue to behave as before, except that their separations have changed because they no longer have identical average spacings between the nodes. Whereas the node spacings are expected to be closer where the particle is defined by its discrete symmetry, their spacings are further apart outside this immediate region. As conjectured earlier, perhaps this node spacing difference in the two regions keeps the lattice total energy value at zero. One now has a lattice with non-uniform node spacings everywhere compared to the original uniform lattice that has no fundamental particles.

Transponders around the electron will continue to scatter the muon's higher frequency feedback signals isotropically into all directions. The lattice distortion will cause these feedback signals to return to the muon out-of-phase with returning feedback signals from other directions, thereby reducing the total amplitude from the forward direction toward the electron, as shown in the Fig. 8 momentum space diagram.

Therefore, the original spherical symmetry of the returning feedback signals around the muon is gone and the muon must either move toward or away from the electron. One can appreciate that the out-of-phase returning signals reduce the total feedback signal amplitude from the electron's direction, which means that the muon will begin to move toward the electron. Why? As shown in Fig. 8, the center-of-momentum for the muon's feedback signal distribution has moved toward the electron. So there is an attraction toward the other particle.

What does the less massive electron do? The same, but in the opposite direction toward the muon of greater mass M. The feedback signals going to the muon region are returned to the electron out-of-phase. Again, the out-of-phase returning feedback signals reduce the total amplitude arriving from the muon's direction, resulting in electron movement toward the muon. That is, the electron's center-of-momentum distribution has moved toward the muon. There is a mutual attraction between the two particles.

The acceleration of each particle occurs when there is a change in phase of the feedback signals arriving from any direction. For example, suppose the particle "senses" that a "chunk" of k-space is absent near k_0-Δk, as shown in Fig. 9. This situation occurs when returning feedback signals from the forward direction are out-of-phase with the oscillation phase of the particle itself. The center of the momentum rectangle will move from k to k' corresponding to a faster moving electron with k' > k, meaning that the particle has moved ahead of the expected uniform velocity location in the corresponding coordinate diagram.

The acceleration is caused by feedback signal amplitude changes as a result of phase changes in the feedback signals as the particle approaches a mass M, an effect directly related to the distortions in the lattice geometry around M. This distortion produces the spacetime curvature associated with GTR gravitation, as explained in the next Section.

In Fig. 10 are shown our two "neutral" particles of masses *m* and M, with *m* approaching the distortion volume around M. One sees immediately for the same angle θ subtended by the feedback signal ray toward M as *m* approaches M, there will be a shorter distance of roundtrip travel for the feedback

signals as they approach one another. And the feedback signals from m will sample regions of greater and greater lattice distortions upon moving closer to the center of M.

In Fig. 11 is an approximation to the result of both effects on the momentum-space amplitude distribution for the two positions shown in Fig. 10, i.e., r_1 and r_5. As more and more amplitude is missing, the change in momentum will increase, i.e., the acceleration toward M will increase upon nearing M as the momentum value increases toward $+k_x$. This type of behavior is expected for the gravitational interaction, because the lattice distortion amount depends upon the mass-energy of M.

The feedback signals are scalar waves given by Eq. 4 in the form $(A/r)\exp[i(\omega(k)t - kx)]$ that are emitted, scattered, and returned, so we can go from the momentum space to coordinate space behavior using the Fourier Transform to obtain the total amplitude at the new, accelerated position for the wave packet

$$U(x) = \int_{k_0-\Delta k_x}^{k_0+\Delta k_x} A(k_0)\ \exp\ [i\,(\omega(k)t - k_x x)]\ dk_x. \qquad (28)$$

And if we assume

$$\omega(k) = \omega(k_0)\ +\ (k - k_0)\left(\frac{d\omega}{dk}\right), \qquad (29)$$

then the composite feedback signal at the electron, i.e., the total amplitude at its new accelerated position is

$$U(x) = \quad 2\Delta k_x\, A(k_0) \exp\,[i\,(\omega(k_0)t - k_0 x)] \quad \frac{\sin \Delta k_x(x - v_0 t)}{\Delta k_x(x - v_0 t)}. \qquad (30)$$

This modulated monochromatic wave does not spread in time, an important property of this feedback signal approach for the behavior of particles.

As $v \to c$, the ellipsoids become more prolate, the Δk_x increases with each equal time interval, and the wave packet of the electron adjusts smoothly. In the limit, the sideward sampling of the environment does not extend beyond the Compton wavelength λ_c and the feedback signals are sampling less of the surrounding space, thus reducing any further acceleration. This behavior agrees with the special theory of relativity (STR).

By considering the acceleration in more detail, one would discover that the smaller range in wave numbers in momentum space spreads the particle wave packet in the x-direction. When a new constant velocity is achieved, the particle wave packet reverts to its normal size. In the perpendicular y- and z-directions in which $v_y = v_z = 0$ as before, a symmetrical hole appears in k_y-space and k_z-space during the acceleration but returns to normal when the acceleration is done. Hence, some temporary lateral spreading of the wave packet occurs also in these directions perpendicular to the accelerated motion along the x-direction.

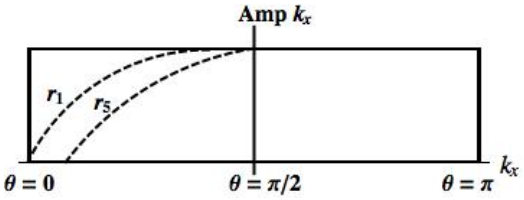

Fig. 11: As the particle approaches M, the feedback signals returning from the two transponder rings have a different phase than for the free particle. The possible reduction of the amplitudes in k-space are shown for positions r_1 and r_5 in Fig. 10, with contributions to the k-space distribution removed above the dashed lines for a range of angles. (approximate representation exagerated).

One could consider further properties of the electron in terms of its de Broglie wavelength = h/p for non-relativistic momentum values in order to discuss the wave packet behavior for the electron. However, the feedback signal approach is all that's needed to understand the electron's behavior in response to another particle that also distorts the lattice.

I have described particle motion in terms of its dependence upon the integral of all the feedback signals returning from the environment back to the source-receiver location. Equal weighting for all k values has been used. In the idealized acceleration example, a rectangularized "chunk" of k-space was missing. Actually, one should consider that some of the feedback signals are returning from all directions with a different phase with respect to the k-space signals returning in a uniformly spaced euclidean lattice. The phase differences would produce a "hole" in k_x-space that can have positive and negative values. All the possibilities could be examined via computer simulations.

13 Gravitation from the Radius Excess

A lattice distortion occurs not only at the particle's origin but also throughout the surrounding space and spacetime. No longer does the lattice have uniformly spaced nodes. As we move further and further away from the origin of each particle, this lattice distortion becomes less and less.

The physics consequences can be understood by first separating the analysis into two parts: the 3-D space part, and then the time part for the (3+1)-D spacetime of our physical world. The two parts are put together to assemble the spacetime of Einstein's GTR.

13.1 The 3-D space part

In the uniformly spaced 3-D sublattice part of C^2 with nodes but with no particles yet, consider an imaginary thin spherical shell with a radius R >> d, the lattice node spacing. Then euclidean geometry dictates a radius value from its surface

area A

$$R = \sqrt{\frac{A}{4\pi}}. \tag{31}$$

Now in this 3-D sublattice consider the electron to have been in existence at the origin so that its characteristic distortion exists everywhere with the amount of distortion decreasing inversely with distance. At the one Compton wavelength distance from the center of the particle, that is, greater than about 10^{-12} meters from the electron's distortion center, one is far enough away to consider an imaginary spherical surface surrounding the electron as a good approximation at all further radial distances.

We measure the distance in a discrete space by counting the nodes along a radial path. Therefore, the measured radius r_{meas} from the electron's center to any outside distance will be greater than for the undistorted lattice because nodes will have been pulled inward. In fact, the radial difference between the distorted lattice and their euclidean lattice is called the *radius excess* expressed by [14]

$$\text{Radius excess} = r_{meas} - \sqrt{\frac{A}{4\pi}}. \tag{32}$$

Note that in the limit when the enclosed mass-energy inside R is reduced to zero, then the radius excess will reduce to the previous zero value. Therefore, let the radius excess be directly proportional to the enclosed mass-energy amount m, in this case the mass of the electron. Then do a dimensional analysis to predict

$$\text{Radius excess} = r_{meas} - \sqrt{\frac{A}{4\pi}} = \frac{G}{3c^2}m. \tag{33}$$

The factor 1/3 comes from the geometry of a 3-D sphere and is the numerical factor for the second term in the Taylor expansion of the sine function.

This radius excess is the important quantity which, according to Einstein's GTR, is indeed proportional to the mass of the particle enclosed by the imaginary sphere at radius R. That is, for a fixed R value, the distance measured by counting the nodes will be greater for the more massive particle enclosed. Note that the radius excess defined here is a measure of the 3-D geometrical curvature produced by the mass-energy m, and that this radius excess expression actually defines the average curvature just above the chosen surface area.

The quantity $G/3c^2 \sim 2.5 \times 10^{-28}$ meters/kilogram, a very small number. Therefore, in order to get a "feeling" for the radius excess magnitude, insert the pertinent values to learn that the radius excess for the electron is extremely small:

$$\text{Electron}: \quad \text{radius excess} = 2.3 \times 10^{-60} \text{ meters!} \tag{34}$$

Also, for Earth: 1.5 millimeters; for the Sun: 0.5 kilometers.

13.2 The time part of (3+1)-D spacetime

Now for the time coordinate contribution. The principle of equivalence states that one cannot distinguish between a gravitational field and an accelerated reference frame for a locally uniform gravitational field. Applying this equivalence principle, Einstein found that time varies from place to place.

The time coordinate will be modified near the mass m. Let v be the relative velocity between a source and a receiver, with the received frequency ω' being related to the emitted frequency ω by Eq. 6 for STR. For $v^2/c^2 << 1$, the approximation is

$$\omega' = \omega\,(1 + v/c). \tag{35}$$

If the receiver is accelerating, then the receiver will have an additional velocity gt, where g is the acceleration value and t is the time interval it takes light to travel the distance H from source to receiver.

Using the equivalence principle, the g is now the gravitational acceleration and H becomes the radial height difference in the gravitational field. For the clock at the radial height h_2 above the clock at height h_1, with $H = h_2 - h_1$,

$$\omega_2 - \omega_1 = \frac{gH}{c^2}, \tag{36}$$

so that the excess rate is

$$\omega_1 \frac{gH}{c^2}. \tag{37}$$

From STR, there is the correction factor of the opposite sign for the speed in case of the moving clocks

$$\omega_2 = \omega_1 \sqrt{1 - v^2/c^2}, \tag{38}$$

which for low speeds $v << c$, becomes

$$\omega_2 = \omega_1(1 - v^2/2c^2), \tag{39}$$

predicting the defect in the rate of the moving clock to be

$$-\omega_1 v^2/2c^2. \tag{40}$$

Combining the two effects produces

$$\Delta\omega = \omega_1 \left(\frac{gH}{c^2} - \frac{v^2}{2c^2}\right). \tag{41}$$

This frequency shift of the moving clock means that if one measures a time interval dt on a fixed clock, the moving clock registers the time interval

$$dt\left[1 + \left(\frac{gH}{c^2} - \frac{v^2}{2c^2}\right)\right]. \tag{42}$$

Therefore, the total time excess over the whole trajectory is the integral

$$\frac{1}{c^2}\int \left(\frac{gH}{c^2} - \frac{v^2}{2c^2}\right) dt, \tag{43}$$

which is to be a maximum, thereby obeying the principle of least action. I.e., the particles always take the longest proper time. Note that this law does not rely upon any of the coordinates.

One can see this result better in the alternative formulation by multiplying Eq. 43 by -mc^2, where m is the mass of the particle, so that the integral is over the kinetic energy minus the gravitational potential energy which, by the principle of least action, must be a minimum.

13.3 The two main laws of GTR

Therefore, the two main laws of GTR have been established by starting from the idea that each fundamental particle distorts the lattice into its own discrete symmetry. The distortion continues to all distances, and phase changes in the returning feedback signals are produced by the distorted lattice.

Equivalently, the distorted lattice around each particle is the source of the radius excess proportional to the enclosed mass producing the distortion, and this radius excess leads to the two main laws of gravitation.

These laws are:

1. The curvature expressed in terms of the excess radius is proportional to the mass inside a sphere, by Eq. 33.
2. Objects move so that their proper time between two end conditions is a maximum.

The first law, Einstein's field equation, reveals exactly how the geometry of spacetime changes in the presence of matter. The second law, Einstein's equation of motion, reveals how objects move when there are only gravitational forces. So the entire spacetime is distorted in the presence of matter.

Can we understand the factor of about 10^{40} for the relative strength of the electric force to the gravitational force between the two electrically charged particles, two electrons, for example. There is a significant physical and conceptual difference between the two forces. The electric force relies upon the local gauge interaction of the SM by the exchange of virtual photons, whereas the gravitational force as determined by the feedback signal approach does not have the exchange of a virtual particle for a local gauge interaction. The gravitational acceleration results from particle responses to their returning feedback signals from the environment. Whether the factor of about 10^{40} can be derived by exploiting this difference is expected but has not been achieved at present.

14 Review of steps taken

Here are the sequence of steps taken to establish that QM and GTR have a common origin determined by the feedback signal approach, based upon the fact that QM, the SM, STR, and GTR are all successful theories that agree with Nature:

1. The lepton and quark particle states respect the electroweak symmetry SU(2) x U(1) of the SM, but the actual two orthogonal fundamental particle states per fermion family are dictated by the discrete symmetry binary subgroups of the unit quaternion group Q, or equivalently, SU(2).
2. The two physical orthogonal EW flavor states in each lepton and quark family are formed by the linear superposition of the two mathematical states, and they oscillate at the Compton frequency ω_C as 3-D entities in R^3. Hadrons combine their 4-D quarks and gluons to make 3-D particles also, obeying QCD.
3. One assumes that (3+1)-D spacetime corresponds to a 2-D complex lattice $C^2 = R^4$ filled with uniformly spaced mathematical nodes acting as ideal transponders.
4. The fundamental fermion "gathers in" the mathematical nodes to form its correct discrete symmetry binary subgroup with its lattice distortion extending outward into the lattice.
5. The "breathing mode" flavor state oscillations of the particle emit scalar waves into the lattice. I have called these "feedback signals".
6. The transponders in the lattice "scatter" these feedback signals into all directions isotropically with no phase shift and with the same response for all frequencies and amplitudes.
7. STR, the principle of inertia, Mach's principle, the path integral approach, QM, and the one direction of time, are all derived by analyzing the details of the feedback signal behavior.
8. The lattice distortion around each fundamental particle is the source of phase changes in the returning feedback signals at the original particle, resulting in an acceleration toward the other particle.
9. Gravitational curvature is shown to agree with the lattice distortion associated with each particle, so the acceleration produced by the feedback signal approach is the gravitational acceleration of GTR.
10. Therefore, QM and GTR have the common origin as established by the behavior of particles in response to the feedback signals.

15 Summary

This feedback signal approach toward understanding particle behavior successfully explains the origin of QM, the path integral method that allows one to calculate quantum mechanical and classical physics behavior, and gravitational acceleration. The approach involves fundamental particles behaving as "antennas" emitting and absorbing scalar waves at their Compton frequencies, scalar waves that I have called feedback signals. These feedback signals are scattered isotropically by a discrete lattice of nodes representing spacetime.

Gravitation has been shown to be the consequence of the lattice distortion around particles by changing the amplitude and phase of the feedback signals that are returning from regions surrounding mass-energy concentrations, in agreement with the radius excess derivation of GTR.

Therefore, I have revealed the common origin for gravitation and quantum mechanics.

The remaining question is whether fundamental particles, such as the electron, do indeed emit and receive these feedback signals as described in this approach. If so, then not only must fundamental particles be using these feedback signals but also all composite entities such as a proton and very massive objects must rely upon them for determining their physical behavior.

Acknowledgements

The author thanks Sciencegems.com for continuing support for fundamental theoretical investigations into problems of importance toward understanding the behavior of Nature.

Submitted on June 27, 2019

References

1. Susskind L. Dear Qubitzers, GR = QM. arXiv: 1708.03040v1 [hep-ph].
2. Tanabashi et al. (Particle Data Group). The Review of Particle Physics, 2018. *Phys. Rev. D98, 030001*.
3. Potter F. A Naive Feedback Model of Particle Motion. *Department of Physics and Astronomy research paper files, University of California, Irvine* , 1982.
4. Potter, F. Discrete internal symmetry groups for leptons and quarks. The Fourth Family of Quarks and Leptons, 2nd International Symposium, proceedings, D.B. Cline & Amarjit, eds. 1989.
5. Potter, F. Geometrical basis for the Standard Model, *Int. J. Theor. Phys.*, 1994, v. 33 279–305.
6. Potter, F. Geometrical Derivation of the Lepton PMNS Matrix Values. *Progress in Physics*, 2013, v. 9(3), 29. Online: www.ptep-online.com/2013/PP-34-09.PDF.
7. Potter F. Exact Neutrino Mixing Angles from Three Subgroups of SU(2) and the Physics Consequences. Workshop in Neutrinos (WIN) 2017 UC Irvine, June 19-24, 2017 Online: indico.fnal.gov/event/9942/session/4/contribution/23/material/slides/0.pdf
8. Potter, F. CKM and PMNS mixing matrices from discrete subgroups of SU(2). *J. Phys.: Conf. Ser.*, 2015, v. 631, 012024. Online: iopscience.iop.org/article/10.1088/1742-6596/631/1/012024/pdf.
9. Baez, J. This Week's Finds in Mathematical Physics (Week 105). Online: math.ucr.edu/home/baez/week105.html, 1995. (accessed 6/6/2019).
10. Feynman, R.P. & Hibbs, A.R. Quantum Mechanics and Path Integrals. Dover Publications, Mineola, NY, 2010.
11. MacKenzie, R. Path Integral Methods and Applications. arXiv: quant-ph/0004090.
12. Feynman, R.P. QED: The Strange Theory of Light and Matter. Princeton University Press, Princeton, NJ, 2014.
13. Feynman, R.P., Leighton, R.B., Sands, M. The Feynman Lectures in Physics: Quantum Mechanics. Addison-Wesley Publishing, Reading, MA, 1965.
14. Feynman, F. Chapter 42: Curved Space. Online: www.feynmanlectures.caltech.edu/11_42.html.
15. Arhrib, A. and Hou, W.-S. CP Violation in Fourth Generation Quark Decays. arXiv: 0908.0901v1 [hep-ph].
16. Hou, George W.-S. Source of CP Violation for the Baryon Asymmetry of the Universe. arXiv: 0810.3396v2 [hep-ph].
17. Wilczek, F. QCD made simple. *Physics Today*, 2000, v. 53(8), 22–28.
18. Pich, A. Quantum Chromodynamics. arXiv: hep-ph/9505231v1.
19. Sheffer, A. Kuratowski's Theorem. Online: www.math.caltech.edu/~2014-15/2term/ma006b/10%20Planar3.pdf (accessed 6/7/2019).
20. Conway, J.H. and Sloane, N.J.A. Sphere Packings, Lattices and Groups, 3^{rd} ed. Springer-Verlag, New York, 1998.
21. Fitzpatrick, R. Two-Electron System. Online: farside.ph.utexas.edu/teaching/389/lectures/node96.html (accessed 6/6/2019).

Appendix A: Quark states

I have proposed [4, 5, 8] that the 4 discrete symmetry binary subgroups that define four 4-D quark families in R^4 are: [333], [433], [343], and [533], corresponding to the only regular polytopes in R^4. The predicted quark mixing angles produce values that generally agree with their empirical values in the standard 3x3 CKM submatrix of its 4x4 quark mixing matrix CKM4. This quark family mixing therefore guarantees that the 4 discrete symmetry binary subgroups defining the quark families collectively behave as the SU(2) of the SM.

Having 4 quark families creates two different conflicts: (1) no 4th quark family has been discovered yet, and (2) there needs to be triangle anomaly cancellation, usually assumed to mean 3 lepton families paired against 3 quark families but with no verification of which lepton family pairs with which quark family. With regard to the first conflict, the mass values of the 4th family quarks could be quite large, so that either they cannot be produced at the LHC [15] or they decay too quickly. The triangle anomaly gets resolved directly because the collective lepton family mimicking SU(2) exactly cancels the collective quark family mimicking SU(2), one-to-one.

The influence of the 4th quark family may yet appear in rare decays of the other quarks and might resolve several extant problems, including being the source of the baryon asymmetry of the Universe (BAU) by providing a needed factor of at least a 10^{13} increase [16] in the Jarlskog constant and by also explaining the muon g-2 discrepancy.

Therefore, the 4-D quark states are clearly distinguished from the 3-D lepton states, the leptons not being capable of having a color charge, which is now a 4-D property. The origin of the three color charge states comes directly from 4-D rotations, which require two simultaneous rotations in orthogonal planes, and there are only three different pairs of orthogonal planes in R^4. The three different color charges, r,g,b, defined by simultaneous rotations in the three pairs of orthogonal planes, can be shown equivalent to the three color charges of SU(3)-color. Even more important, having quark states and gluon states defined in R^4 means they cannot ex-

ist in R^3, so quark confinement becomes geometrically explained also.

Finally, the 4-D quark and gluon states must combine according to quantum chromodynamics (QCD) to make the mesons and baryons, i.e., the 3-D hadrons. Intersection theory in mathematics can handle this geometrical concept of intersecting 4-D objects to make 3-D objects.

However, QCD theory predicts [17, 18] a self-contained world for the quarks and gluons, with only color changes allowed and no possibility of quark decay. So why does Nature need the leptons? The mathematical answer follows from Kuratowski's theorem [19] in graph theory: all graphs will reduce to the K_5 or $K_{3,3}$ graphs, the only graphs that retain their integrity. Fortunately, at least for quarks, the [333] discrete symmetry binary subgroup of the up/down quark family represents the K_5 graph, so all other quarks will decay eventually to this first family. The stability of the electron may also be a consequence because [332] is related to [333].

Also recall that only 4N-dimensional normal spaces have a conjugate space of the same dimension according to Clifford algebra and Bott periodicity [9]. So, there will be the simultaneous existence of the 4-D anti-particle real internal symmetry space as required by the SM. The next larger space with a conjugate space, R^8, is equivalent to a 10-D spacetime. For discrete spaces, icosians related to the binary icosaheral group [532] provide a direct connection [20] from our discrete R^4 to the discrete space R^8, which obeys the discrete symmetry operations of Weyl E_8.

The particles exist in our discrete SO(3,1) spacetime, so the icosians produce a second discrete symmetry Weyl E_8 for spacetime. Combining discrete spacetime with the discretee internal symmetry group therefore makes the discrete product group Weyl E_8 x Weyl E_8, equivalent to the discrete symmetry group I call "discrete" SO(9,1). Hence, there exists a *unique* connection from the SM gauge group to "discrete" SO(9,1) in a 10-D spacetime.

Appendix B: Identical particles and quantum statistics

Consider two identical particles. What behavior will the feedback signal approach predict?

Two neutral identical particles are to be considered, so that we can ignore any local gauge interactions of the SM, both particles beginning at rest with respect to each other. In the general case, feedback signals emitted at the same Compton frequency ω_1 by each particle are absorbed, phase shifted, and emitted by the other identical particle back into the surrounding space.

Their existence in each other's environment means that the identical particles can become phase-locked, either with in-phase or with out-of-phase normal modes, as is the case for two identical-frequency quantum harmonic oscillators communicating to each other, with their final locked-in phase relationship becoming 0 or π.

The two possible normal mode frequencies for any two harmonic oscillators communicating via an exchange of energy represented by Γ are

$$\Omega = \frac{1}{2}(\omega_1 + \omega_2) \pm \Gamma, \tag{44}$$

but the two identical high Q fundamental particles will have $\omega_2 = \omega_1$, so

$$\Omega = \omega_1 \pm \Gamma. \tag{45}$$

Which physical property of a particle actually determines the difference between the two phase-locked states? Because the single free particle does not have phase-shifted returning feedback signals, the phase shifts introduced by the other identical particle can be a function of differences only:

$$\text{phase shift} = f(\omega_i - \omega_j, A_i - A_j, P_i - P_j), \tag{46}$$

where ω is the Compton frequency, A is the signal amplitude, and the P could be some other factor such as the intrinsic spin.

As we know, the physical factor P called particle intrinsic spin S is the key. Different particle angular momentum spin states need to be considered, such as a scalar S = 0, a spinor S = 1/2, and a vector S = 1, in order to determine the general result.

Consider the scalar particles first, the ones with intrinsic spin S = 0. At first the feedback signals returning from the direction of the other identical scalar particle might not be in-phase, so the two particles are accelerated toward each another because the returning feedback signals from the vacuum transponders in the direction opposite the other particle are in-phase. Eventually, the scalar particles can become locked in-phase with each other's oscillations and can occupy the same point in space. So these two S = 0 identical particles behave as bosons obeying Bose-Einstein statistics.

Now consider a system of two spin S = 1/2 electrons. QM requires [21] that their overall asymmetric wavefunction be the product of position eigenvalues and the total spin quantum numbers. There are three triplet spinor states having S = 1 symmetric with respect to the exchange of the electrons, with the spatial part being asymmetric so that the probability of the two electrons being at the same point in space is zero. But for the singlet S = 0 spinor state, the spin part is asymmetric and the spatial part is symmetric, thereby enhancing the probability to be at the same point in space, i.e., there is an attraction to one another.

Applying geometry by rotating the two S = 1/2 identical particles together in the triplet S = 1 state by 360°, one determines that the feedback signals will return with a phase that produces an increased amplitude pushing each particle away from the direction of the other identical particle. Therefore, a repulsion occurs to produce an increased separation. Called Pauli repulsion, this response is the source of Fermi-Dirac statistics.

In the total S = 0 case for two spin S = 1/2 particles, i.e., with spins opposite, the feedback signal amplitudes at each particle decrease by adding in the returning signals from the direction of the other identical particle. There is attraction, so this total S = 0 spin state is allowed for two electrons at the same point in space. That is, the spatial wavefunction is even but the spin wavefunction for this total S = 0 state is anti-symmetric.

Finally, when both particles each have S = 1, the total spin states are S = 2 and S = 0. The geometrical factors will produce a result identical to the total S = 0 Bose-Einstein behavior for two scalar particles, i.e., there is a feedback signal amplitude decrease that results in an attraction.

Back to Cosmos

F. M. Sanchez[1], V. A. Kotov[2], M. Grosmann[3], D. Weigel[4], R. Veysseyre[5], C. Bizouard[6], N. Flawisky[7], D. Gayral[8], L. Gueroult[9]

[1]Professor (retired), Université Paris 11, Orsay, France. E-mail: hol137@yahoo.com.
[2]Astronomer, CrAO*, Nauchny, Crimea, Russian Federation. E-mail: vkotov43@mail.ru.
[3]Professor (retired), Université Louis Pasteur, Strasbourg, France. E-mail: michelgrosmann@me.com.
[4]Professor (retired), Université de Dijon et Paris 6, France. E-mail: dominiqueweigel118@gmail.com.
[5]Professor (retired), École Centrale, Paris, France. E-mail: renee.veysseyre@gmail.com.
[6]Astronomer, OBSPM, Paris, France. E-mail: christian.bizouard@obspm.fr.
[7]Architect/Math Instructor, ENSA† Paris Malaquais, France. E-mail: flawisky@free.fr.
[8]Computer Science Engineer. E-mail: denis.gayral@epita.fr.
[9]Lecturer (retired), ENSA Paris Malaquais, France. E-mail: lgueroult@hotmail.com.

The antique concept of a permanent Cosmos is reintroduced as a perfect deterministic computer, inverting the Anthropic Principle and interpreting the dimensionless parameters as optimal calculation bases. The later are unified in the Topological Axis, which exhibits the string theory dimension series $d = 4k + 2$, with the emphasis on the values 26 (visible universe) and 10 (the hydrogen-pion couple). The 1-D extension of the Holographic Principle defines the Grandcosmos and a 10^{61} trans-plankian quantified time. This confirms the matter-antimatter oscillatory bounce and resolves at last the vacuum energy dilemma. The intervention of the sporadic groups implies the mathematics-physics fusion which is confirmed by 10^{-9} precise relations, showing four force connection with the Eddington constant 137 and the Atiyah one. The Holic Principle, the generalized Holographic Principle and Eddington's theory must unlock particle physics, with composite d quark and massive string, gluon, photon and graviton. The standard evolutionary cosmology will soon be excluded by the observation of mature galaxies in the very far-field.

Contents

1 The hierarchy and computation principles

There is presently an intense debate in the physics community. While a minority believes in an Ultimate Theory, a large majority have abandoned such hope and believes seriously in the extreme consequence of the "Anthropic Principle", the Multiverse conundrum [1]. The present article settles the debate in favor of a *single steady-state flickering* cosmos (Section 4), a kind of synthesis between the two historic main cosmologies, since it can be viewed as a *Permanent Big Bang*.

Only a minority thinks physics and mathematics are really unified, while a large majority separate the two domains (so separating also biology). The criteria for the uniqueness of the Cosmos is the mathematical character of the measured dimensionless parameters. Indeed, we show in Section 2 that *the latter obeys the Topological Axis, Fig. 1, and, for the first time, they are connected with a series of ppb relations involving e, π and γ* (Section 9.4). This article shows also that the discovery of the sporadic groups, with, in particular, the monstruous moonshine correlation [2], is a crucial discovery for

*Crimean Astrophysical Observatory
†École Nationale Supérieure d'Architecture

physics (Section 8.5).

In this debate unicity-multiplicity, pure mathematicians believe that progress can be obtained only when the Ultimate Theory has been discovered. However, the history of physics shows that one can progress without knowing the ultimate laws. This no-said principle can be called the *"Hierarchy Principle"*. So, when Proust and Dalton found whole numbers in chemical reactions, they were prefiguring atomic physics. The same for Balmer, spectral lines and wave mechanics. Idem for Mandeleev, atomic masses and nuclear physics. Also, when Mandel found whole numbers in biology, he anticipated genetics. In the same manner, *this article prefigures the fundamental theory, but precising its arithmetical foundation: the Holic Principle*, recalled in Section 6. We interpret this central role of whole numbers by assuming that the Cosmos is a perfect computer. This is the very foundation of quantum physics. The Section 3 shows the overall holographic quantification, breaking the Planck wall by a factor 10^{61}, solving at last the vacuum quantum energy dilemma and *justifying why the Cosmos is so large*. This *"Optimal Computation Principle"* enlightens the First Principle of Thermodynamics, the energy conservation. This is a more direct and logical explanation than the standard "time uniformity".

This reinstates the Laplace determinism, involving non-local hidden variables, which are identified with the Cosmos, so rejecting the standard Copenhagen statistical interpretation of quantum mechanics. It seems that the pre-scientific role of chance is a common point between three misleading views in present mainstream thinking. Firstly, in biology, the assimilation of Darwin's rough argumentation with a scientific theory (see Discussion). Secondly, in quantum physics, the so-called "uncertainty principles", which are only manifestations of the general wave propagation (field and *flickering* matter), through Fourier transform properties. Thirdly, in cosmology, the above recurse to the Multiverse conundrum.

While it was already shown that main dimensionless parameters are present both in musical scales and in DNA characteristics [3], this article goes further, by showing they are *calculation bases*.

The abnormal efficiency of elementary 3-fold dimensional analysis is justified in Section 5, confirming the reality of the Grandcosmos, essential in Coherent Cosmology [3]. The c-free analysis gives simply and directly the supercycle period in an all-deterministic Cosmos, with dimension $d = 30$, given by the Holic Principle. An elementary calculation gives also a good approximation of the *invariant* Hubble radius, in a formula which was present for a century in astrophysics textbooks: the limit of a star radius when the number of atoms reduces to unity. We recall that in Coherent Cosmology, the Hubble radius R is defined by the relative redshift law

$$\Delta f/f = l/R$$

of l-distant galaxy groups, in the *exponential* recession.

Finally, there is the central problem of infinity. While it is welcome in mathematics, it is condemned in physics. The domination of mathematics blocked for years the quantum mechanics, annonced by the above discoverers, from Proust to Mandel. Indeed, Planck believed in the *mathematical continuum*, and was reluctant of his own *physical discovery*, until 1912, when Poincaré demonstrated that the quantification of matter-light interaction was mandatory [4]. The continuum has the advantage that it simplifies formulas, by the virtue of the computation properties of e and π. Thus, the vastness of the Cosmos is a compromise, but at the expense of *a necessary rationalization of e and* π, as shown in this article.

Thus, there must exist multi-base algorithms able to explain the compatibility between these two principles, Hierarchy and Computation, which seems at first sight somewhat contradictory. The key is the analysis of the dimensionless parameters (about 30 in the standard model), which are tightly contrived by a mysterious "fine-tuning". Happily, the Hierarchy Principle applies: only three dimensionless parameters: a, p, and a_G are sufficient to explain the main structures of the world [1]. Two of them are precisely measured: the electric constant $a \approx 137.035999139(31)$, known with 0.23 ppb precision, and the proton-electron mass ratio $p \approx 1836.15267245$ (75), known with 0.4 ppb precision. The gravitational coupling constant a_G was the square of the ratio Planck/proton mass, subjected to a relatively large imprecision 10^{-4} due to the imprecision on G measurement. In fact, we consider rather the inverse of α and α_G, we note a and a_G.

One reads [1]:

> For example, the size of a planet is the geometric mean of the size of the Universe and the size of an atom; the mass of man is the geometric mean of the mass of a planet and the mass of a proton. Such relationships, as well as the basic dependencies on α and α_G from which they derive, might be regarded as coincidences if one does not appreciate that they can be deduced from known physical theory, with the exception of the Universe, which cannot be explained directly from known physics... This line of arguments, which is discussed later, appeals to the 'anthropic principle'.

This is misleading since, as soon as the fine-tuning involves the observable Universe radius, it signals the existence of a fundamental theory that must take into account the *antique Cosmos concept, which, as Eddington claimed [5], must be permanent*. Extending this to the standard spatial homogeneity, this leads to the Perfect Cosmological Principle, the very foundation of the steady-state cosmology and the starting point of Coherent Cosmology [3].

2 The cosmic fine-tuning and the topological axis

We look here for a systematic organization of dimensionless physical quantities stemming from cosmology, astrophysics, particle physics, theoretical physics and mathematics. The most famous fine tuning implies cosmic quantities, awkward-

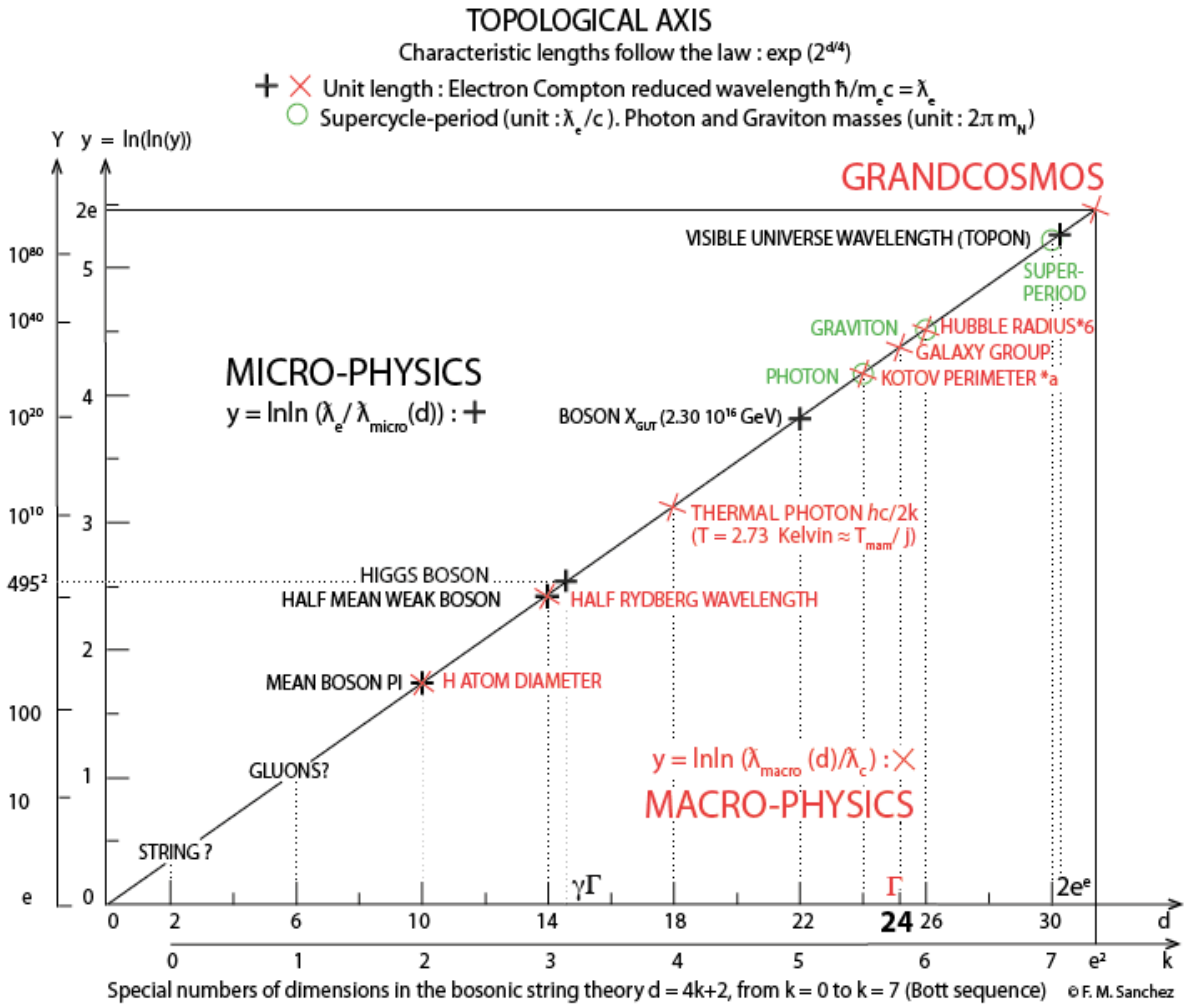

Fig. 1: The Topological Axis (data in Table 1). The double natural logarithms $y = \ln(\ln(Y))$ of the main dimensionless physical quantities Y corresponds to the special string dimension series $d = 4k + 2$, from $k = 0$ to $k = 7$, characteristics of the Bott sequence [27]. This is the reunion of height 2D-1D holographic relations, hence the name "Topological Axis". Two relations come from the double large number correlation [5], one comes from the Carr and Rees weak boson-gravitation relation (2), and one comes from the Davies analysis [11], involving the Cosmological Microwave Background (CMB) wavelength. *On the macrophysics side, with length unit $\lambda\!\!\!^{-}_e$, the electron Compton reduced wavelength*, 6 × Hubble radius 13.812 billion light-years, (3), is tied to the bosonic critical dimension 26, while Bott reduction $\Delta d = 8$ leads firstly to $d = 18$: it is the *thermal photon* (CMB). This temperature $T \approx 2.725820805$ Kelvin, (38), is identified to the common temperature of the couple Universe-Grandcosmos. It is tied to the the mammal wavelength through the Sternheimer scale factor j (Section 8.3); another Bott reduction leads to $d = 10$ (superstring dimension): it is the *hydrogen atom*, and finally to $d = 2$: the *massive string*, about 2.1 GeV. For the number 24 of transverse dimensions, it is the *Kotov length* (Section 4.3), multiplied by a factor about $2\pi a$, with $a \approx 137.036$. For $d \approx \Gamma$, the Atiyah constant (Section 8.2), it is the *galaxy group* radius, a characteristic cosmic length (10^6 light-years, Section 2.1). For $k \approx e^2$, $y \approx 2e$, it is the *Grandcosmos* radius (Section 3). The Space-Time-Matter Holic dimension $d = 30$ (Section 6) is tied to c times the cosmic *supercycle period* (Section 5). *On the microphysics side, with the same length unit $\lambda\!\!\!^{-}_e$*, Bott reductions from $d = 30$ lead to the *gauge bosons*: $d = 22$ for the Grand Unification Theory (GUT) one, (2.30×10^{16} GeV), $d = 14$ for the weak one and $d = 6$ for the (*massive*) gluons, about 8.6 MeV. For the intermediary superstring value $d = 10$, there is the mean *pion*. For $d \approx \gamma \times \Gamma$, $Y \approx 495^2$ the square of the diminished Green-Schwarz string dimension (496 − 1), it is the *Brout-Englert-Higgs boson* (125.175 GeV). For $k \approx 2e^e$, it is the *topon*, the visible Universe wavelength, the space quantum, which identifies with the monoradial unit length of the Bekenstein-Hawking Universe entropy (Section 3). *With unit 2π times the Nambu mass $m_N = am_e$* [15], $d = 24$ and 26 corresponds to the *photon and graviton masses*, defined by the two-step holographic interaction [3], Section 7.1. This is the extrapolation towards smaller numbers of the Double Larger Number correlation. The central dimension is $d = 16$, for a total of 2^7 string dimensions in the Bott sequence. This suggests a liaison with the Eddigton's matrix 16 × 16 [5].

Table 1: Topological Axis $f(d) = \exp(2^{d/4})$. Data with $R = 2a_G\lambda\!\!\!^{-}_e = 2\hbar^2/Gm_em_pm_H \approx 13.812$ Glyr, $R_{GC} = 2r_e^6/l_P^5 \approx 9.0758 \times 10^{86}$ m

$\lambda\!\!\!^{-}_M = 2l_P^2/R \approx 3.9989 \times 10^{-96}$ m, $T = \hbar^4/\rho_c^{3/2}G_F^{5/2} \approx 5.4829 \times 10^{57}$ s, $l_K = \lambda\!\!\!^{-}_e(a_Ga_w)^{1/2}$, $m_{ph} = a_wm_{gr} \approx 1.222 \times 10^{-55}$ kg

Physical element	k	$d = 4k + 2$	$\ln(\ln(f(d)))$	$\ln(\ln($Measured ratio$))$ [17]	Predictions ($\lambda\!\!\!^{-}_e = ar_e = ct_e$)
string	0	2	0.347		$m_{string} \approx 2.1$ MeV ?
gluon	1	6	1.040		$m_{gluon} \approx 8.6$ MeV ?
mean pion	2	10	1.733	$\ln(\ln(268.60)) \approx 1.722$	
H atom diameter	2	10	1.733	$\ln(\ln(274.22)) \approx 1.725$	
half mean weak boson	3	14	2.426	$\ln(\ln(8.378 \times 10^4)) \approx 2.428$	
Higgs boson	-	$\gamma\Gamma \approx 14.533$	2.518	$\ln(\ln(2.449 \times 10^5)) \approx 2.518$	$m_{Higgs} \approx 125.175\ GeV$?
thermal photon	4	18	3.119	$\ln(\ln(hc/2k\theta_{CMB}\lambda\!\!\!^{-}_e)) \approx 3.035$	
boson GUT	5	22	3.812		$m_{GUT} \approx 2.30 \times 10^{16}$ GeV ?
photon	5.5	24	4.159		$\ln(\ln(m_N/m_{ph})) \approx 4.130$
Kotov perimeter	5.5	24	4.159	$\ln(\ln(2\pi l_K/r_e)) \approx 4.159$	
Hubble radius R*6	6	26	4.5054	4.506(3) [6]	$\ln(\ln(6R/\lambda\!\!\!^{-}_e)) \approx 4.5054$
graviton	6	26	4.505		$\ln(\ln(m_N/m_{gr})) \approx 4.485$
supercycle period	7	30	5.199		$\ln(\ln(T/t_e)) \approx 5.199$
topon	-	$2e^e$	5.253		$\ln(\ln(\lambda\!\!\!^{-}_e/\lambda\!\!\!^{-}_M)) \approx 5.523$
Grandcosmos	e^2	-	5.432		$\ln(\ln(R_{GC}/\lambda\!\!\!^{-}_e)) \approx 5.433$

ly called the "Double Large Number Problem". If it is a "problem" for standard evolutionary cosmology, it is a precious clue in the steady-state cosmology based on the above *Perfect* Cosmological Principle (spatial *and* temporal homogeneity). This cosmological fine-tuning leads directly to a *gravitational hydrogen molecule model of the visible universe* [3].

This defines the Universe Hubble radius $R = 2a_G\lambda\!\!\!^{-}_e$, where the factor 2 comes from the bi-atomic structure, and where $\lambda\!\!\!^{-}_e = \hbar/cm_e$ is the electron Compton reduced wavelength, while the gravitational coupling constant is $a_G = \hbar c/Gm_pm_H$, where m_p and m_H are the proton and hydrogen atom masses. So, *the speed c is eliminated*, in accordance with the Coherent Cosmology which needs signal celerity far exceeding c. This gives $R \approx 13,812$ Gly, corresponding to a Hubble constant 70.790 (km/s)/Megaparsec, compatible with the most recent measurements [6]: 72(3) (km/s)/Megaparsec. The latter confirms the value measured by the 1a type novae, while the standard optimization of 6 parameters results in a lower value, by 9%. This is a significant refutation of the standard cosmology, but the fact that the so-called Universe age is about 13.8 Gyr cannot be due to chance. This means that the standard approach has something right [10], but the standard interpretation is false: in fact the Big Bang is permanent.

Consider the wavelength of the visible Universe with critical mass $M = Rc^2/2G$:

$$\lambda\!\!\!^{-}_M = \hbar/Mc \approx 4.00 \times 10^{-96}\,\mathrm{m}\,. \tag{1}$$

This "topon" corresponds to the value $n \approx 2e^e$, close to the touchstone $n = 30$ of the Topological Axis, see Fig. 1. This scheme illustrates the function $f(n + 4) = f^2(n)$ and stems from the imbrication of relations of the form $\lambda\!\!\!^{-}_e/l_{micro} \sim (l_{macro}/\lambda\!\!\!^{-}_e)^2$, followed by $l_{macro}/\lambda\!\!\!^{-}_e \sim (\lambda\!\!\!^{-}_e/l'_{micro})^2$, leading to:

$$\lambda\!\!\!^{-}_e/\lambda\!\!\!^{-}_M \sim (R/\lambda\!\!\!^{-}_e)^2 \sim (\lambda\!\!\!^{-}_e/\lambda\!\!\!^{-}_X)^4$$
$$\sim (\lambda_{CMB}/\lambda\!\!\!^{-}_e)^8 \sim (\lambda\!\!\!^{-}_e/\lambda\!\!\!^{-}_W)^{16} \sim (2r_H/\lambda\!\!\!^{-}_e)^{32}$$
$$\sim (\lambda\!\!\!^{-}_e/l_{Gl})^{64} \sim (\lambda\!\!\!^{-}_{str}/\lambda\!\!\!^{-}_e)^{128} \sim 2^{2^8}\,.$$

This series include the Cosmic Microwave Background wavelength λ_{CMB} and a string wavelength $\lambda\!\!\!^{-}_{str}$, with mass about 2 MeV. Hence, the correlation is eight-fold. They include implicitly the above double fine-tuning and three more relations that have been independently reported [3]. Thus, only three relations are really new. The overall large number 2^{256} has an obvious computational character, confirmed below by the dramatic appearance of the Eddington Large Number.

In particular, as Davies quoted [11] *"The fact that* $R/\lambda_{CMB} \sim a_G^{3/4}$ *seems to indicate yet another large-number coincidence"*. By this order of magnitude, we infer rather precise relations. With the hydrogen radius r_H, we observe $R/r_H \approx (4\pi\lambda_{CMB}/r_H)^4$, precise to 0.6%. Considering the standard cosmological neutrino background (CNB), which wavelength is defined by $(\lambda_{CNB}/\lambda_{CMB})^3 = 11/4$, we note that $R/\lambda\!\!\!^{-}_e \approx (\lambda^2_{CNB}/\lambda_{CMB}\lambda\!\!\!^{-}_e)^4$ to 1.7%. The appearance of the neutrino field is conform with the synthesis of the two main cosmologies, where the single Bang is replaced by a matter-antimatter Oscillatory Bounce [10].

It was noted in [1] that a_G is of order W^8, where W is the W boson-electron mass ratio. With the above R value, one observes the following more symmetrical relation involving the other (neutral) weak boson Z, in the 0.01% indetermination of W and Z:

$$R/(\lambda\!\!\!^{-}_p\lambda\!\!\!^{-}_H)^{1/2} \approx (WZ)^4 \tag{2}$$

where $\lambda\!\!\!^{-}_p$ and $\lambda\!\!\!^{-}_H$ are the proton and hydrogen reduced wavelengths. The precision of this formula will be pulled to the ppb range in Section 9.4, by intervention of canonical mathematical constants.

The gravitational hydrogen molecule model [3] implies the following double correlation, which is the simplest case of Eddington's statistical theory [5]: the position of a "reference particle" is supposed to be determined with an uncertainty of $R/2$. For N particles of mass m components of the visible Universe, the deviance is statistically divided by $\sqrt{N}$, where $N = M/m$. If m is the principal value of the effective mass of the electron in the hydrogen atom, $m = m'_e = m_e m_p/m_H$, and if, moreover, one equates the deviance $R/(2\sqrt{(M/m'_e)})$ to the hydrogen reduced wavelength $\lambda\!\!\!^{-}_H = \hbar/cm_H$, one gets:

$$R/2\lambda\!\!\!^{-}_H = (M/m'_e)^{1/2} = \hbar c/Gm_e m_p\,. \tag{3}$$

This is the definitive interpretation of the Double Large Number fine-tuning. So, while the two pillars of physics, relativity and quantum theory are unable to conciliate gravitation and particle physics, the third pillar, statistical physics, directly makes this connection in cosmology [5].

Recall that, contrary to what is often stated, quantum physics does not limit to microphysics. Indeed, the exclusion principle applies in both solid state physics and in stellar physics. In particular, for a star containing N_s atoms, in which the pressure has reached the quantum degeneracy value (case of white dwarfs), exclusion principle applies for electrons, and the star radius is about $R/N_s^{1/3}$ [3]. So the formula giving the Hubble radius R, a very difficult measurement which puzzled a whole century, was implicitly contained in astrophysics textbooks. Eddington was aware of this Cosmologic Exclusion Principle, but he could not conclude since, at his epoch, the Hubble measurement for R was false by an order of magnitude.

The reason for this discrepancy is that *Lemaître and Hubble considered galaxies of the Local Group, which do not participate in the so-called space expansion*. In fact, it is sufficient to introduce a repulsive force proportional to separation distance, for explaining the steady-state exponential recession. *The repulsive force is equivalent to reintroduce the Einstein cosmological constant in the General Relativity equations, but with invariant value* $1/R^2$.

The distance for which this force exceeds attractive gravitation between galaxies is about 10^6 light years [3], a typical galaxy group radius, which corresponds, in the Topological Axis, to the Atiyah constant Γ, (Section 8.3), see Fig. 1.

In the steady-state cosmology of Bondi, Gold [7] and Hoyle [8], such a repulsive force between galaxy groups is necessary, in order to avoid a big chill due to the thermodynamics second principle. But, inside a galaxy group, another evacuation mechanism must occur: *it would be the role of the massive black holes*.

3 The toponic holographic quantification

In the above steady-state cosmological model, the Perfect Cosmological Principle implies the invariance of the Universe mean mass density ρ, defined at large. This predicts also the exponential recession of galaxy groups, with time constant R/c being compensated by the appearance of m_n massive neutrons at rate c^3/Gm_n, corresponding to about one neutron by century in a cathedral volume. The invariant visible Universe radius R is then defined by the Schwarzschild relation, so that each topon, with wavelength $\lambda\!\!\!^{-}_M = \hbar/Mc = 2l_P^2/R$ is the center of an equivalent R-radius black hole, of critical mass $M = Rc^2/2G$. The Bekenstein-Hawking entropy of this black hole Universe shows a 1-D extension [3] of the standard Holographic Principle, until now devoted to 3-D application only [12]:

$$S_{BH} = A/4 = \pi(R/l_P)^2 = 2\pi R/\lambda\!\!\!^{-}_M \tag{4}$$

where A is the horizon sphere area and $l_P = (G\hbar/c^3)^{1/2}$ is the Planck length. Note that, while the standard evolutionary cosmology uses differential equations, which are not adapted to a single Universe, as Poincaré stated [9], the Permanent Cosmology must favor such integral relations. Here it is the *Archimedes testimony tying the disk area to its perimeter*.

The topon breaks the so-called "Planck wall" by a factor $l_P/\lambda\!\!\!^{-}_M \approx 10^{61}$. This explains why this holographic relation was long time unnoticed. Indeed, it was admitted that l_P was the quantum of space: in fact *the Planck length is an intermediate holographic length only*.

The gravitational potential energy of a critical homogeneous sphere is $-(3/5)GM^2/R = -(3/10)Mc^2$, while the *non-relativistic* kinetic energy of galaxies is $(3/10)Mc^2$ [3]. Their sum is therefore zero: the density of the so-called "dark energy" is compatible with 7/10, so that dark energy was a trivial false problem. The relativity theory is a local theory that does not apply in cosmology at large: galaxies actually reach speed c, and, crossing the horizon, enter a Grandcosmos of radius R_{GC}, given, as a first approximation, by the symmetrical monochrome holographic relation:

$$S_{BH} = \pi(R/l_P)^2 = 2\pi R^{(0)}_{GC}/l_P \tag{5}$$

with $R^{(0)}_{GC}/R = l_P/\lambda\!\!\!^{-}_M \sim 10^{61}$. The conservation of the time constant $t = R/c = R^{(0)}_{GC}/C$ introduces a canonical velocity $C \sim 10^{61}c$, lifting the veil on an energy larger than that of the visible Universe by a factor of 10^{122}, which can be identified with the l_P-normalized quantum energy of vacuum, checked by the Casimir effect [13]. *The central problem of quantum cosmic physics is thus solved*. Moreover, the objections against the Hawking approach using transplankian frequencies are wiped out [14].

In a better approximation, justified below, R is replaced in the above relation by $R' = 2\hbar^2/Gm_N^3 \approx 18.105$ Gly, where $m_N = am_e$ is the Nambu mass [15], of central importance in

particle physics. Indeed, the half radius $R'/2$ has a simpler definition than $R/2$: it corresponds to the elimination of c between the classical electron radius and the Planck length [3]. In this way, the sphere of radius R' appears as the spherical hologram representation of the outer Grandcosmos:

$$S'_{BH} = \pi(R'/l_P)^2 = 2\pi R_{GC}/l_P \,. \tag{6}$$

This value will be confirmed in Section 5 (Fig. 6).

The toponic quantification hypothesis assumes that the mass of a particle is an exact sub-multiple of the critical mass M of the visible Universe: $m = M/N_m$. Thus its wavelength is $N_m \lambda\!\!\!^{-}_M$, allowing the following holographic extension of the above monoradial holographic conservation:

$$S_{BH} = \pi(R/l_P)^2 = 2\pi R/\lambda\!\!\!^{-}_M = 2\pi N_m R/\lambda\!\!\!^{-}_m \,. \tag{7}$$

This series of diametrical circles generate, by scanning, the approximation of a sphere: thus it goes from the disk to the sphere with area $4\pi(R/l_P)^2$. Note that *this justifies the factor* $\frac{1}{4}$ *in the BH entropy*. But, for the approximation to be sufficient, *the numbers* N_m *must be very large*. In this way, the Cosmos computer can use the computational properties of the mathematical constants of the continuous analysis, such as e and π, (Sections 8 and 9).

The immensity of the Cosmos thus receives a computational holographic explanation, which is much simpler than that of standard cosmology, where initial conditions, during Planck time, would be adjusted with extreme precision, even with inflation.

With $N_{Ed} = 136 \times 2^{256}$ the Eddington large number, one observes that N_{Ed} times the neutron mass, corrected by the classical ratio H/p, gives the effective mass $3M/10$ to 41 ppm, so that:

$$Mm_p = m_P^4/m_e m_H \approx (10N_{Ed}/3)m_H m_n \tag{8}$$

This directly involves the Planck mass m_P, which presently has no known interpretation, except that it is close to the mass of the human ovocyte [3]. In this way, the local inertia is related to the distant masses, in accordance with the Mach principle, which the relativity theory does not explain. Another shortcoming of this theory is that it does not define any inertial frame. However, the Doppler asymmetry of the cosmic background indicates that the speed of our local group of galaxies is about 630 km/s. The cosmic background is, therefore, tied to the Newton absolute frame, the Grandcosmos.

The mathematical continuity is excluded by the above Computation Principle, so the time associated to the above "topon":

$$t_M = \lambda\!\!\!^{-}_M/c = \hbar/Mc^2 \approx 1.33 \times 10^{-104} s \tag{9}$$

is the new candidate for the "chronon", the "quantum of time", so the oscillatory bounce has a frequency about 10^{104} Hz [10]. The CPT symmetry (Charge conjugation-Parity inversion-Time reversal) connects this matter-antimatter oscillation with the parity violation in particle physics and biology.

4 The tachyonic flickering space-time-matter

The tachyonic hypothesis is consistent with the non-local character of quantum mechanics.

4.1 The single electron cosmology

The single-electron cosmology [3] uses the electron indeterminacy, which is the real basis of the Exclusion Principle, giving a horizon value R_1 only dependent of the principal value of the hydrogen radius $a' = aH/p$, by respect to $\lambda\!\!\!^{-}_e$. It is the value for which the mean cosmic value is also the atomic one:

$$\frac{\sum(1/n)}{\sum(1/n^2)} = a' \tag{10}$$

with the sum running from 2 to $R_1/\lambda\!\!\!^{-}_e$. This implies:

$$R_1 = \lambda\!\!\!^{-}_e \exp((\pi^2/6 - 1)a' + 1 - \gamma) \approx 15.77465\,\text{Gly}$$

very close (0.4 ppm) to $R_1 = (p_G/p_0)(\beta R R')^{1/2}$, where $p_G = P/2^{127/2}$, with $P = \lambda\!\!\!^{-}_e/l_P$, $\beta = (H - p)^{-1}$ the Rydbergh correction factor and $p_0 = 6\pi^5$ the Lenz-Wyler value p (Section 9.2). Moreover, there is a direct connection with the Grandcosmos radius and the topon, to 0.90 %:

$$\lambda\!\!\!^{-}_M = 2l_p^2/R \approx R_1^3/R_{GC}^2 \,. \tag{11}$$

This *synthesis relation* confirms the coherence of the whole procedure. It will be of central importance in the following.

4.2 The Cosmic Coherent Oscillation (CCO)

The Kotov non-doppler cosmic oscillation [16] is not considered seriously, since it seems to violate the most basic prerequisite of physics, the generality of Doppler phenomena. Interpreting this as a tachyonic phenomenon, we identified the Kotov period $t_K \simeq 9600.06(2)$ s, taking the electron characteristic time $t_e = \lambda\!\!\!^{-}_e/c$ as unit, to the simplest relation eliminating c between a_G and $a_w = \hbar^3/G_F m_e^2 c$, the well measured (3×10^{-7}) dimensionless electroweak coupling constant a_w :

$$t_K/t_e = (a_G a_w)^{1/2} \,. \tag{12}$$

This weak coupling constant [1] $a_w = (E_F/m_e c^2)^2$ is defined from the Fermi energy [17]: $E_F \approx 292.806161(6)\,\text{GeV} \approx 573007.33(25)\, m_e c^2$, itself tied to the weak force constant $G_F \equiv (\hbar c)^3/E_F^2 \approx 1.4358509(7) \times 10^{-62}\text{Joule} \times \text{m}^3$. This introduces the product of two area speeds, confirming the flickering hypothesis:

$$(\lambda\!\!\!^{-2}_e/t_K)(\hbar/(m_p m_H)^{1/2} = (GG_F)^{1/2} \tag{13}$$

so the best measured cosmic quantity, the Kotov period, implies a symmetry between gravitation and weak nuclear force.

This specifies the G value to 10^{-6} precision (ppm). It is compatible with the well-elaborate (10^{-5}) BIPM measurement [18], at several sigmas from the Codata value [17], but the later is the mean between discordant measurements. Computer analysis shows that this value of G is compatible with the following well-defined value, with $d_e \approx 1.001159652$ the relative electron magnetic moment [17] :

$$(2^{127}/a_G)^{1/2} \approx d_e(H/p)^3$$
$$\Leftrightarrow$$
$$G \approx 6.6754552 \times 10^{-11}\,\mathrm{kg^{-1}m^3s^{-2}}.$$

A value ppb confirmed in Section 9. One notes:

$$\sqrt{(R_1/a_w l_K)} \approx 4\pi p/p_0 \leftrightarrow t_K \approx 9600.591445\,\mathrm{s}$$

a relation independent from G. This Kotov period t_K value will be confirmed, in the ppb range, in Section 9.4. It is associated [3] with the photon mass $m_{ph} = \hbar/c^2 t_K \approx 1.222\times10^{-55}$ kg. The connection with the graviton mass is proposed in Section 7.1.

The following relation (0.1%), will be very useful in the Section 5:

$$M/m_{ph} \approx (3/e)O_M^2 \qquad (14)$$

with O_M the cardinal order of the Monster group [19]. The Monster Group, the largest of 26 sporadic groups, is suspected by some researchers to play a central role in physics: indeed string theory allows a bridge between apparently unconnected mathematical theories [2].

4.3 The omnipresence of CCO in astrophysics

With $t = R/c$, the relation $(t\,t_K^2)^{1/3} \approx 10.8$ years, compatible with the famous 11-year sun period was noted. It was proposed that this unexplained phenomenon, responsible for moderate periodic climate variation, was also of flickering cosmic origin [20]. This hypothesis has been recently confirmed by the straight temporal profile of the phenomena, showing it is tied to a quantum process [21].

Remarkable enough, a "mysterious" period $\approx 1/9$ days of the Sun's pulsations has been predicted long before its actual discovery in 1974. Namely, 73 years ago, French amateur astronomer Sevin (1946) claimed that "la période propre de vibration du Soleil, c'est-a-dire la période de son infra-son (1/9 de jour), a joué un role essentiel dans la distribution des planètes supérieures". Presumably, the Sevin "vibration period" of the Sun was merely an issue of his reflections about resonances and distances inside the solar system. Nevertheless, solar pulsations with exactly that period were discovered, after decades – and independently of Sevin's paper – by a few groups of astrophysicists. Soon the presence of the same period, or timescale, was found in other objects of the Cosmos too [16].

Opponents emphasize often that t_K is very close to the 9$^{\text{th}}$ harmonic of the mean terrestrial day: the corresponding ratio – of the length of a day to the t_K period – is equal to 8.99943(1) – and claim thus the t_K oscillation of the Sun should be regarded as an artifact (see, *e.g.* Grec and Fossat, 1979; Fossat *et al.*, 2017). As a matter of fact, however, the t_K period occurs to be the best commensurate timescale for the spin rates of all the most massive and fast-rotating bodies of the solar system, in general.

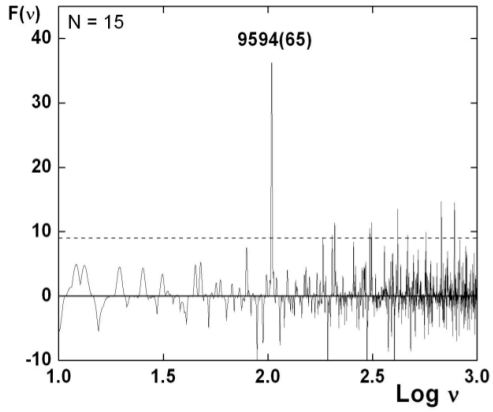

Fig. 2: Resonance-spectrum $F(\nu)$ computed for 15 motions of the largest, fast-spinning bodies of the solar system. On horizontal axis is logarithm of frequency ν in μHz, the dashed horizontal line shows a 3θ C.L., and the primary peak yields to the best – commensurable period 9594(65) s.

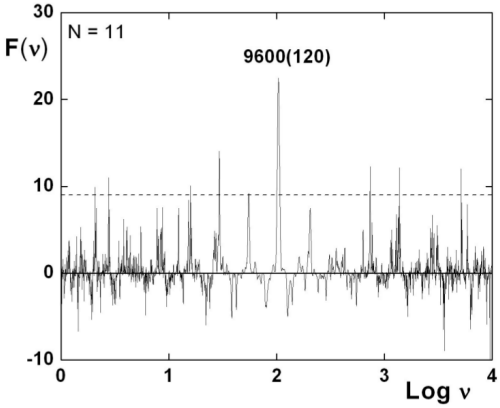

Fig. 3: Same as Fig. 2, for $N = 11$ sizes "diameters" of the solar system (with $c = 1$ and the π factor for inner orbits). The highest peak corresponds to the spatial scale 9600(120) light-sec.

This is obvious from Fig. 2, which shows the resonance spectrum $F(\nu)$, calculated for 15 motions of 12 largest, fast spinning, objects of the system (with the mean diameters $\geq$ 500 km and periods inferior to 2 days: six planets, three asteroids and three satellites, leaving apart trans-neptunian objects; see Kotov, 2018). The peak of the best commensurability corresponds to a period of 9594(65) s, which coincides well, within the error limits, with t_K at about 5.3θ C.L., *i.e.*

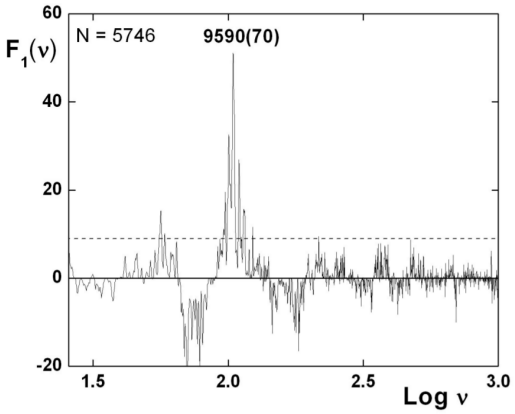

Fig. 4: Resonance-spectrum $F_1(\nu)$, computed for $N = 5746$ binaries with periods inferior to 5 days. Horizontal axis gives logarithm of the trial frequency ν in μHz, the dashed line indicates a 3θ C.L., and the major peak corresponds to a timescale of 9590(70) s.

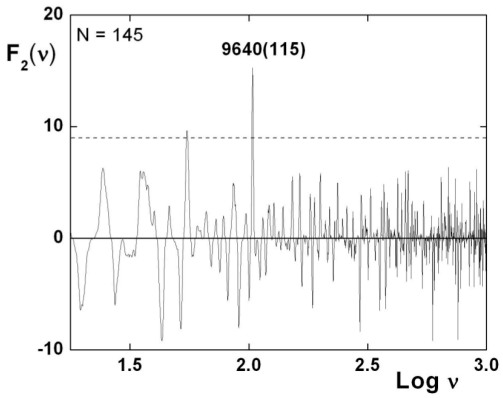

Fig. 5: Same as Fig. 4, for the $F_2(\nu)$ spectrum, computed for $N = 145$ exoplanets with P inferior to 1.5 days. The strongest peak of the composite commensurability corresponds to a period of 9640(115) s, at nearly 3.9θ significance (after Kotov, 2018).

with a chance probability 10^{-7}.

It seems very puzzling also that the spatial scale $l_K \approx 19.24$ A.U. occurs to be the best commensurate with orbital sizes of the main planetary orbits of the solar system, – see Fig. 3, where the resonance spectrum $F(\nu)$ is plotted for 11 orbits, including those of asteroid belt, Pluto and Eris (orbital "diameters" were approximated by the major axes, and for the inner orbits they were multiplied by π). The primary peak – of the best commensurability – corresponds to the spatial scale 9600(120) light-sec., or 19.24(3) A.U., at 4.7θ C.L. (Kotov, 2013).

Close binaries are characterized by the t_K resonance too, with the π number as a factor of ideal incommensurability of motions, or frequencies (Kotov, 2018). Fig. 4 shows the resonance spectrum, or metrics of motion, $F_1(\nu) \equiv F(\pi \times \nu/2)$, computed for 5746 close binaries, including cataclysmic variables and related objects. The major peak, with C.L. of about 7θ, corresponds to the timescale 9590(70) s, coinciding within the error limits with t_K (the stellar data were taken from all available binary stars catalogues and original papers).

To compute the $F_1(\nu)$ spectrum, the program finds – for each test frequency ν – deviations of ratios $(2\nu_i/\pi\nu)k \geq 1$ from the nearest integers, and determines then the least-squares minimum of such deviations. Here, ν is the test frequency, ν_i minus the frequency of a given object, $i = 1, 2, ...N$ – the ordinal number, with N, the total number of observed periods in a sample of objects, and the power $k = 1$ or -1. The factor of two in Eq. (2) takes into account that second half of the orbit repeats the first one, and the transcendental number π appears as a factor of orbital stability, or "idea" incommensurability, of motions, or frequencies (the π number, in fact, characterizes geometry of space; for details see Kotov, 2018).

Recently it was shown, that the t_K timescale characterizes, statistically, the motion of superfast exoplanets too, see Fig. 5.

It was shown in fact, that a number of superfast exoplanets, with periods inferior to 2 days, revolve around parent stars with periods, near-commensurate with timescales t_1 and /or $2t_1/\pi$, where $t_1 = 9603(85)$ s agrees fairly well with the period $t_K \approx 9600$ s of the so-called "cosmic oscillation" found firstly in the Sun, then – in other variable objects of the Universe (the probability that the two timescales would coincide by chance is near 3×10^{-4}).

4.4 The Tifft, Arp and Pioneer effects

Another unexplained effect is the 75(5) km/s periodicity in the galactic redshift [22]. Now, this speed $v_1 \approx ca/F$ corresponds to the following quantum resonance, with the electron classical radius $r_e = \lambda\!\!\!^{-}_e/a$ and where $m_F = m_e\sqrt{a_w}$ is the Fermi mass:

$$v_n/n = v_1 = \hbar/r_e m_F \,. \tag{15}$$

The Halton Arp observations of chains of galaxies with different redshifts [23] was also rejected. But it could be the sign of the galactic regeneration constantly maintaining the visible Universe mass: this is sustained by the following section proving the invariance of the mean mass density ρ_c.

Much controversial is the Pioneer deceleration [24] $g_{Pi} \approx 8.7 \times 10^{-10}\,\mathrm{ms}^{-2}$. It corresponds to the Pioneer time $t_{Pi} = c/g_{Pi} \approx 3.4 \times 10^{17}$ s close to $t = R/c \approx 4.3587 \times 10^{17}$ s. The following section will show a connection between the Kotov, Tifft and Pioneer effects.

5 The logic of prospective dimensional analysis

Physics uses principally *physical quantities* of the type $Q = M^x L^y T^t$, where M, L and T are Mass, Length and Time measurements, and where the exponents are rational numbers. However, the addition of measures of different categories has no significance. This seems at first sight illogical since, fundamentally, a product is a sum of additions. So, *there must*

be a hidden common nature for the three categories, mass, length and time. This sustains the above single electron cosmic model [3].

This suggests a 3-D geometrical model. Indeed, consider $t = R/c$, and $M' = R'c^2/2G$ the critical mass in the above holographic sphere representing the Grandcosmos. Summing the square of $\ln(M'/m_e)$, and two times the square of $\ln(R/\lambda\!\!\!^{-}_e) = \ln(t/t_e)$, one gets, to 40 ppm:

$$\ln^2(M'/m_e) + \ln^2(R/\lambda\!\!\!^{-}_e) + \ln^2(t/t_e) \approx \ln^2(R_{GC}/\lambda\!\!\!^{-}_e) \qquad (16)$$

showing the Grandcosmos ratio. This traduces, in function of $P = m_P/m_e$, $p = m_p/m_e$, $H = m_H/m_e$ by:

$$\ln^2(P^4/a^3) + 2\ln^2(P^2/pH) \approx \ln^2(2P^5/a^6)\,. \qquad (17)$$

Moreover, to 10^{-7}, corresponding to 7×10^{-6} precision on the above G value:

$$\ln^2(P^4/a^3) + 2\ln^2(P^2/pH) \approx \exp(4e{-}1/a)\,. \qquad (18)$$

This is a dramatic geometrical confirmation (Fig. 6) of the visible Universe-Grandcosmos holographic couple.

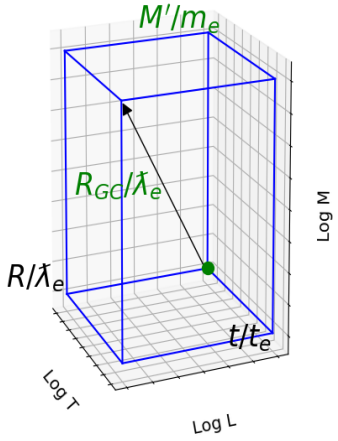

Fig. 6: *Geodimensional Universe-Grandcosmos couple*, with unit length the electron Compton reduced wavelength. In a 3-D superspace, logarithms of physical ratios are considered vectors. The Grandcosmos radius appears as the norm of the vector using for length and time projections the same value $R/\lambda\!\!\!^{-}_e = t/t_e$. For the mass projection it is M'/m_e where M' is the critical mass in the Grandcosmos reduced spherical hologram. This is a dramatic geometrical confirmation (not dependant of the base for logarithms) of the Extended (2D-1D) Holographic Principle applied to the Bekenstein-Hawking Universe entropy (6). The Grandcosmos existence cannot be denied since the relation involving *natural* logarithms with e and a reach precision 10^{-7}.

Another crucial point in physics is the existence of invariant fundamental constants. Thus, association of three of them must give characteristic values of M, L, T. So, approaching a domain in physics necessitates to calculate characteristic values (M, L, T) from the three universal constants which are the most pertinent in the considered domain. This prospective dimensional analysis is largely used in fluid mechanics, where the equations are intractable. However, it is largely ignored in other domains because there is not really mathematical foundation, apart the above essential remarks. The triplet $c, G, \hbar$ which define the above Planck units is a notable exception.

Moreover, in virtue of the above Hierarchy Principle, the lack of theoretical justification is not a reason to neglect prospective dimensional analysis.

The elimination of c in the above R formula means that the simplest basic dimensional analysis starting from $\hbar$, G and m, the electron-proton-neutron mean mass, gives a good approximation for $R/2$. Indeed, in the hypothesis of a coherent Cosmos, it is logical to discard c which is far two small a speed. This has not been observed during one century since c is always believed to be the single mandatory foundation of space-time. The warning of Poincaré [25], the true discoverer of relativity: *"use 4-D space-time, but do not confound Space and Time"* has long been forgotten, and physicists have unwisely put $c = 1$ in their equations.

In his three first minutes of cosmology (Sept. 1997), the first author obtained the length:

$$l\{\hbar, G, m\} = \hbar^2/Gm^3 \approx R/2 \qquad (19)$$

but it took nine years to get this published [20], and it appeared later [3] that m must be considered more precisely as the cubic root of the product $m_e m_p m_H$. Moreover, the above critical condition links the time $t = R/c$ and the mean mass density by the c-free formula:

$$\rho_c = 3/8\pi G t^2 \approx 9.41198\times10^{-27}\,\mathrm{kg}\times\mathrm{m}^{-3}\,. \qquad (20)$$

Thus, the mainstream idea of a temporal variability of the mean density ρ_c cannot be to sustain, meaning that ρ_c must be considered a fundamental constant. This writes:

$$t\{\hbar, \rho_c, G\} = 1/\rho_c^{1/2}G^{1/2} = (R/c)(8\pi/3)^{1/2}\,. \qquad (21)$$

This idea of ρ_c being a fundamental constant permits to define the Hubble radius R without any ambiguity: this is the radius of the sphere containing a critical mass. This justifies the above application of the Bekenstein-Hawking entropy.

Opponents would say that the center of a black hole presents a singularity: that is indeed the case for the topon in the above flickering space- mass-time hypothesis. Others will argue that the flying galaxies cannot reach the celerity c at horizon, but, as recalled above, relativity is a local theory, so do not apply to cosmology at large. Indeed, even General Relativity in unable to define any Galilean frame, while the Foucault pendulum shows it directly, realizing the Cosmic Microwave Background frame, identified with the Grandcosmos frame, as seen above.

Introducing the Fermi constant G_F, the associated c-free length is very particular, to 1.7%:

$$l\{\hbar, \rho_c, G_F\} = \hbar/\rho_c^{1/2}G_F^{1/2} \approx 9.07154\times10^9 m \approx \lambda\!\!\!^{-2}_e/l_P\,.$$

Now, most dramatically, the following mandatory c-free times are close to each over (0.7%):

$$T\{\hbar, \rho_c, G_F\} = \hbar^4/\rho_c^{3/2} G_F^{5/2} \approx 5.4829 \times 10^{57}\,\mathrm{s} \qquad (22)$$

$$T'\{\hbar, G, m\} = \hbar^3/G^2 m^5 \approx 5.5224 \times 10^{57}\,\mathrm{s}\,. \qquad (23)$$

One would conceive it is the deterministic supercycle period, which matches the Topological Axis at $n = 30$, the holic dimension (see Section 6), to 4%. Comparing T with the Kotov non-doppler Cosmic Oscillation period $t_K \approx 9600.60(2)$ s, one observes, to 0.04 % and 0.2 %:

$$T/t_K \approx O_M/\sqrt{2} \approx e^a/\sqrt{a_w}$$

where O_M, the cardinal order of the Monster Group, have been detected in the Section 4.2, again in relation with the Kotov period. Eliminating the latter, this introduces the above t_M chronon :

$$T/t_M \approx (3/e\sqrt{2}) O_M^3 \qquad (24)$$

The simplest interpretation follows: *this is the number of quantum events in a supercycle of period T, in a perfectly deterministic Cosmos*.

Introducing the above Pioneer abnormal deceleration g_{Pn}, one gets the time: $t\{G, m_e, g_{Pn}\} = (Gm_e/g_{Pn}^3)^{1/4} = (t_{Pn}^3 t'_e)^{1/4}$, where $t_{Pn} = c/g_{Pn}$ and $t'_e = Gm_e/c^3$. This time is compatible with: $t\{G, m_e, g_{Pn}\} = t_K/(F/a)^2$, where the above Tifft factor F/a appears. The implication of the time $t'_e = Gm_e/c^3 = 2.2568 \times 10^{-66}$ s confirms the above Planck wall breakdown.

6 The arithmetical logic: Holic Principle

In the hypothesis of an arithmetical Cosmos, the ultimate equations must be diophantine. The simplest one is $T^2 = L^3$, where T is a time ratio and L a length one, resolved, since 2 and 3 are co-prime, by:

$$T^2 = L^3 = n^6 \qquad (25)$$

where n is a whole number, showing the classical 6-D phase-space of point mechanics. Considering the exponents, this particularizes the usual 3-D space, but attributes 2 dimensions for the time, in conformity with an independent study [26].

This is the degenerate arithmetic form of the 2D-3D holographic principle.

This is *also* Kepler's third law. It was the simplest one of his three laws, and the realization of his research of harmony. Indeed, its diophantine form says more: it gives $L = n^2$, *the orbit law in the hydrogen atom and in our gravitational molecule model*, where the visible Universe corresponds only to the first orbital. This suggests at once the existence of a Grandcosmos.

Before the superperiod was recognized, the first version of the Topological axis [3] showed an overall dissymmetry. This was another sign for the Grandcosmos existence. Now, this corresponds to $d = 30$, the natural extension of the above diophantine equation:

$$T^2 = L^3 = M^5 = n^{30} \qquad (26)$$

where M is a mass ratio. Recall that the lifetime of an unstable particle depends on the 5th power of its mass. This holic dimension 30 is the touchstone of the Topological axis, from which the gauge bosons are deduced by Bott reductions [27] (Fig. 1).

This is called the Holic Principle, but limited to *the apparent MLT world only*. The Complete Holic Principle [29] involves a field term F^7, and so introduces the dimension $30 \times 7 = 210$. This is confirmed by (to 0.56 %, -0.65 %, -0.59 %, -0.32 %):

$$R/\lambda\!\!\!^{-}_e \approx s_4^5 \approx f(26)/6 \approx \Gamma^{28}/5 \approx (2/\delta)^{210} \qquad (27)$$

where $s_4 = 2\pi^2 a^3$ is the area of the 4-sphere of radius a and Γ is the Atiyah constant (Section 8.3). Moreover (0.1 %, 0.03 % and 0.9 %):

$$2/\delta = 2R/R' \approx \ln p/\ln a \approx \ln a/\ln \Gamma \approx \ln \Gamma/\ln f \qquad (28)$$

where f is the inverse strong coupling constant (Section 8.3). This confirms the central computational role of $\delta = R'/R = pH/a^3$, which is to 1.6 ppm: $\delta \approx e^{2/e^2}$. This implies a geo-combinatorial relation between a and p:

$$p^{(p^2)} \sim (a^2)^{(a^3)} \qquad (29)$$

showing a symmetry between basic powers of a and p.

7 The special holographic relations

The holographic technique, based on the properties of a coherent wave, is by far the most efficient way to treat huge information, in particular in optics [28].

The students of the first author realized in 1987 a hologram by scanning a 1 mW security power laser beam upon a photosensitive area of 0.6 m^2. The emulsion depth 10 microns permitted false color to be obtained by varying illumination through a photomask, and use of a shrinkable emulsion chemical process. The information contained in this hologram reached 10^{15} bit, obtained in 12 minutes of scanning exposition. Then, the first author claimed *"such an efficient way of dealing information must be used by Nature"*. Turning to the impressive data of particle physics, after an intensive study, holographic relations were indeed found, and its arithmetical form, the Holic Principle was presented at ANPA 16 (Cambridge, 1994) [29].

In Sept. 1997, the Orsay University attributes a sabbatical year, giving time to reexamine the foundation of cosmology. *In the three first minutes, the half-radius of visible Universe was obtained*. After several weeks, the scanning holography of Section 3 was established. After rejection by the Orsay

University and the French Academy, this was put in March 1998 in a closed draft in the Académie des Sciences de Paris, under the title *"L'Univers conserve-t-il l'information ?"*. The next year, the initial form of the Topological Axis was rejected by the French Academy, when an anonymous referee argued that "le Big Bang est avéré".

Strangely enough, when the first author's publication was blocked (1993-1995), a Holographic Principle was coined by some theoreticians [12], which were not specialists in holography. The origin of this appellation is not clear. One may think that the name comes from the idea of dimension reduction, from 3-D to 2-D, similar to the visual impression in current visible holograms (in fact holography is only the 2-D restitution of a propagating wave). It this respect, it is strange that no one tried to extend this process to 1-D. The idea of temporal 1-D holography was proposed in the first author's thesis as soon as 1975 [30].

While the standard Holographic Principle is limited to using the Plank area, it is natural to suppose that there are other holographic units. In fact, *the Topological Axis is the reunion of eight 1D-2D holographic relations*. We present here four more confirmations.

7.1 The graviton and photon masses

The electromagnetic interaction is not really understood, especially the photon concept [31]. The main lesson of modern physics is that everything (light *and* matter) propagates by waves (quanta appearing only at the detection). This implies directly the non-local hidden variable (Cosmos), without involving the so-called EPR paradox [32]. Indeed, a coherent wave is represented by a unitary operator: we have shown that the quantum formalism is very similar to the holographic one, describing an interaction by a two-step holographic process. We recall that convergent and divergent waves lead to an oscillation [3]. This is known as the particle exchange of a massive boson associated with any interaction. Here, it is assumed that the boson has a tachyonic speed C_1. Now, the resonance condition is that the wavelengths are identical, by analogy with the Gabor condition [33]. So, for the electron wave and the weak wave:

$$\lambda\!\!\!^{-}_e = \hbar/m_e c = \hbar/m_{gr} C_{gr} \tag{30}$$

$$\lambda\!\!\!^{-}_w = \hbar/m_w c = \hbar/m_{ph} C_{ph}\,. \tag{31}$$

Equaling the tachyonic celerities to C_1, and m_{ph} with $\hbar/cl_K$ [34], taking account of the *ppb correlation* tying R_1 and l_K (Section 4.2), one gets:

$$C_1/c = m_e/m_{gr} = m_w/m_{ph} = l_K/\lambda\!\!\!^{-}_w \approx R_1/\lambda\!\!\!^{-}_e(4\pi p/p_0)^2\,. \tag{32}$$

This leads to $m_{ph} \approx 1.2222 \times 10^{-55}$ kg, and $m_{gr} \approx 3.7223 \times 10^{-67}$ kg, which fit canonical places in the Topological Axis (Fig. 1). This means (0.8 %):

$$a_w = m_{ph}/m_{gr} \approx f(26)/f(24) \tag{33}$$

calling for further study.

7.2 The conservation of information

The Grandcosmos holographic reduction radius R' shows in itself an holographic relation with the CMB Wien wavelength $l_{CMB} = hc/kTv$, with $v = 5(1 - e^{-v}) \approx 4.965114245$, and the proton radius, identified, as a first approximation, to $\lambda\!\!\!^{-}_e/\sqrt{D} \approx 8.7029 \times 10^{-16}$ m (0.1 %, -0.1 % and 87 ppm):

$$e^a \approx 4\pi(R'/l_{CMB})^2 \approx (2\pi/3)(r_p/l_P)^3 \approx \sqrt{3} M_B/m_P \tag{34}$$

where $M_B = 2M/\sqrt{n}$ is the baryonic mass of the Universe [3]. The factor $\sqrt{3}$ implies a new holographic relation (see the "neutron relation" in Section 8.3):

$$4\pi(R/e^a l_P)^2 \approx (4\pi/3)n \approx (4\pi/3)(v\pi^2/4)^3\,. \tag{35}$$

Since the holographic technique uses coherent radiation, this seems incompatible with the CMB thermal character. But *in a totally deterministic cosmos, there is no paradox*. This question is connected with the black hole information paradigm [35]. Independently of our approach, an argument in favor of a total conservation of information was tied to a non-evolution cosmology [36].

So, while General Relativity and quantum physics disagree about the nature of space-time, especially the non-locality phenomena, they agree for complete determinism, leading to *the definitive rejection of the Copenhagen statistical interpretation*.

The Wien wavelength enters (0.03 %):

$$l_{CMB}/\lambda\!\!\!^{-}_e \approx P/pHa^3 \tag{36}$$

confirming that the cosmic temperature is invariant. Note that the measured proton "charge radius" $8.775(5) \times 10^{-16}$ m is slightly distinct from the above value. There is presently a "proton radius puzzle" [37].

7.3 The cosmic temperature

In the gravitational hydrogen molecule model [3], the Hubble radius R shows the following 1D-2D special holographic relation, using the wavelengths of the electron, proton and hydrogen, while the background wavelength appears in the logical extension, the 3-D term involving the molecular hydrogen wavelength:

$$2\pi R/\lambda\!\!\!^{-}_e = 4\pi\lambda\!\!\!^{-}_p\lambda\!\!\!^{-}_H/l_P^2 \simeq (4\pi/3)(\lambda\!\!\!^{-}_{CMB}/\lambda\!\!\!^{-}_{H2})^3\,. \tag{37}$$

The above relation gives $T_{CMB} \approx 2.73K$. Moreover it is another dramatic example of c-free dimensional analysis [3]. With the measured temperature of the cosmic background, there is a small gap compatible with $(H/p_G)^2 p/6\pi^5$, where $p_G^2 = P^2/2^{127}$, with $P = \lambda\!\!\!^{-}_e/l_P$. This eliminates l_P, producing a relation independent of G, implying $T_{CMB} \approx 2.725820805$ Kelvin. Recall that $2^{127} - 1$ is the most famous prime number in the history of mathematics, being the last term of the

Combinatorial Hierarchy [3] of special imbricated Mersenne numbers 3, 7, 127, the sum of which is 137 (Section 8.2):

$$2^{2^{2^{2^2-1}-1}-1}-1 = 2\pi^2\lambda_{CMB}^3/\lambda\!\!\!^{-}_e\lambda\!\!\!^{-2}_H \tag{38}$$

which is the area of the 4-sphere of radius $\lambda_{CMB}/\lambda\!\!\!^{-}_m$, where $\lambda\!\!\!^{-}_m = (\lambda\!\!\!^{-}_e\lambda\!\!\!^{-2}_H)^{1/3}$. This proves the relevance of the Lenz-Wyler approximation for the proton/electron mass ratio $p_0 = 6\pi^5$, (Section 9.2).

7.4 The Holic Principle and CCO

The sphere of radius $R' = 2r_e^3/l_P^2$, where $r_e = \lambda\!\!\!^{-}_e/a$ is the electron classical radius is the Grandcosmos hologram (Section 3). Its HB entropy writes: $\pi(R'/l_P)^2 = (\pi/2)(R'/r_e)^3$, *i.e.* with a wrong geometric coefficient. However, the HB entropy of the visible Universe shows a nearly geometric term, with imprecision $4/3\delta \approx 1.017$:

$$\pi(R/l_P)^2 \approx (2\pi/3)(R/r_e)^3 \tag{39}$$

which is a holographic conservation in *the half-sphere of the visible universe*. By analogy with the above scanning process filling the whole sphere (Section 3), the above Kotov length l_k (Section 4.3) permits to introduce two holographic relations, involving the whole sphere (0.90 % and 2.6 %):

$$\pi(R/l_K)^2 \approx 2\pi l_K/r_e \tag{40}$$

$$(4\pi/3)(R/l_K)^3 \approx 4\pi(r_e/l_P)^2\,. \tag{41}$$

The deviation of the first relation is very close to that of (11): $R_1^3 \approx R_{GC}^2\lambda\!\!\!^{-}_M$. This induces, with precision 17 ppm, identified to 0.3 ppm with $np^2/H^2\sqrt{pH}$, and 0.08 %:

$$(R_{GC}l_P/Rr_e)^2 \approx (R_1l_K/Rr_e)^3 \approx 3^{1/3}\mu^{35} \tag{42}$$

showing a quasi-holic form implying μ, the muon/electron mass ratio. The complete holic form with dimension 210 is shown by the study of the BH entropy of the Grandcosmos: (12 ppm, 100 ppm, 42 ppm):

$$\begin{aligned}\mu^{210} &\approx \pi^{-3/2}(R_{GC}/l_P)^4 \approx 4\pi\tau^{137}(a/137)^2\\ &\approx O_M^9 \ln D(p/n)^2\,.\end{aligned} \tag{43}$$

Ths is a perfect illustration of the Hierarchy Principle. Thus the expected correlation [38] [39] of $\ln D$ with 4π is confirmed. The existence of a final theory based on the Holic Principle (Section 6) and the Grandcosmos cannot be denied. The interpretation is clear: *the 4-D space-time of Grandcosmos is associated with a 9-D space involving the Monster. This opens a path towards the Final Theory.*

The term $R_1l_K/R'r_e$ is close to (1%) $\sqrt{O_M} \approx 2a^2P$ (0.18 %). The study of deviations shows the intermediate bosons ratios W and Z, with values specified to the ppb range in Section 9.4, leading to (-4 and 3.5 ppm, 0.3%):

$$O_M(FR'/PR_1)^2 \approx W^4(137/a)^3 \tag{44}$$

$$(F^3R_1/2a^3R')^2 \approx Z^4(a/137)(p_0/p)^2\,. \tag{45}$$

This refines the relation $a_w/WZ \approx \sqrt{a}$ known (0.1 %) in particle theory (0.3 and -0.4 ppm)

$$137p_0W^2Z^2/pa_w^2 \approx \sqrt{O_M}/(2a^2P) \approx e^{1/-4a}\,. \tag{46}$$

Thus, in first approximation ($e^{-1/4a} \approx 0.036$ %), the square root of the Monster order is the ratio of the Rydberg wavelength $2a^2\lambda\!\!\!^{-}_e$ to the Planck length.

8 The role of intermediary mathematical constants

8.1 The electrical constant a

The electrical constant a characterizes the Coulomb force between two l-distant elementary charges at rest:

$$F_{qq} = \hbar c/al^2\,. \tag{47}$$

Since any electrical charge is a whole multiple of unitary charge q (a relativistic invariant), *any electrical force depends only on the above constants and whole numbers*. Hence, it is logical that a appears central in atomic physics and in many fine-tuning relations [1].

However, theorists focused on one property only, the appearance of its fifth power in the hydrogen hyperfine spectra, calling its inverse α, the "fine-structure constant".

Many researchers looked for the mathematical origin of a. In quantum electrodynamics, $\sqrt{a}$ is connected with the electron magnetic abnormal factor, which is very precisely measured [17]: $d_e \approx 1.00115965218076(27)$. It is readily seen that $\sqrt{a} \approx \exp(\pi/2)^2$. From $i = e^{(i\pi/2)}$, this writes $i^{-\ln i}$ and the study of deviation leads to, with $a_e = a/d_e$ (29 ppb):

$$i^{-\ln i}/\sqrt{d_e} \approx (\sqrt{a_e} + 1/\sqrt{a_e})^2\,. \tag{48}$$

The slight deviation is not a valid objection, since *Nature must use rational approximations for* π. Indeed, the fractional development for the corresponding π value is 3, 7, 15, 1, $(\tau/\mu)^2$, with μ and τ the normalised masses of the heavy leptons. It is a formal rationalisation, focussing on an acute problem of present standard model, which is unable to explain the three families of particles. Thus, the study of the muon and tau mass ratios is crucial. One observes (1 ppm, 56 ppm, 0.02 %):

$$\begin{aligned}2/\delta &\approx (1/2d_e)\ln(pH)/\ln a\\ &\approx (1/d_e^2)\ln\tau/\ln\mu \approx d_e^2\ln s/\ln\tau\end{aligned} \tag{49}$$

where s is the Higgs ratio (Section 9.4). The following Koide relation [40], *which has a mathematical justification in terms of circulant matrix* [46], correctly predicted τ at an epoch (around 2000) during which its measurement was false to 3 σ. It writes:

$$(1+\mu+\tau)/2 = (1+\sqrt{\mu}+\sqrt{\tau})^2/3 = p_K\,. \tag{50}$$

This Koide relation, quite discarded by the scientific community, is another sign of the serious incompleteness of the present particle physics standard. This Koide-Sanchez constant will be precised to ppb precision in Section 9.4.

8.2 The Eddington constant 137

The initial Eddington proposal for a was the whole number 136, being the number of independent parameters in the symmetric matrix 16×16. Note that *$n = 16$ is the central dimension of the Topological Axis*. Later, one unity was added, becoming 137 [5]. It shows a symmetry between the 11 dimensions of M theory (a synthesis of five string theories) and the 4 of space-time. Indeed: $137 = 11^2 + 4^2$, while, as seen above: $11/4 = (\theta_{CMB}/\theta_{CNB})^3$.

Since Riemann series are tied to the prime number distribution, it seems odd and incredible that mathematicians have not point out the primes appearing in the harmonic series since it is the single Riemann pole. It seems that the basic precept *all occurs in the pole* was forgotten in this case.

As ancient Egyptians used only fractions of type $1/n$, they were certainly aware of this particular harmonic series: $S_5 = 137/60$. Indeed it appears in the Ptolemaic approximation for π: $\pi_{Pt} = 377/120 = 2 + S_5/2$.

It is strange that Eddington's theory was rejected as soon as a appeared to deviate from 137. Indeed, the following shows that *137 plays a central role in ppb fine-tuning analysis*. Note that Nambu [15] showed that the mass $m_N = 137m_e$ is central in particle physics.

One may interpret $137+1$ as the sum of the numbers of dimensions in the Topological Axis [3], taking into account the double point (H atom-pion couple) for the superstring value $d = 10$, and the remarkable sum:

$$\sum_{k=7}^{k=0}(4k+2) = 2^7 \,. \tag{51}$$

So $137 = 2^7 - 1 + 3 + 7$, *i.e.* the Combinatorial Hierarchy form [41]. But this appears also as $137 = 135 + 2$, showing the string dimension 2. Indeed, one obtains the value $a \approx 137.035999119$ compatible with measurement value in:

$$\ln 137/\ln(a/137) \approx (2 + 135/d_e)^2 \tag{52}$$

meaning that *the ratio $a/137$ acts as a canonical ratio*.

Considering the product of the T.A. dimensions:

$$P_d = \Pi_{k=7}^{k=0}(4k+2) = 2^8 3^4 5^2 7^1 11^1 13^1 \tag{53}$$

which is a simple sub-multiple of the cardinal order of the Suzuki group, and a simple multiple of the three other sporadic groups M_{11}, M_{12} and J_2 [19]. With l_W the mean of the CMB and CNB Wien lengths (0.06 %):

$$P_d \approx l_W/\lambda_e \,. \tag{54}$$

The pertinence of the Topological Axis series is thus confirmed, calling for further study.

8.3 The Atiyah and Sternheimer constants

Sir Michael Atiyah was a precursor in the search for unity in mathematics and physics. In his last work [42], the Bernoulli function $x/(1 - e^{-x})$ plays a central role. *This is the kernel of the thermal Planck law*. Considering the above Wien reduced constant $v = hc/kT\lambda_{Wien}$, one notes that $a \approx e^v - 2\pi$, suggesting a to be a trigonometric line. Indeed $\cos a \approx 1/e$, and, to 65 ppb:

$$a \approx 44\pi - \arccos(1/e) \tag{55}$$

a formula diffused on the web, but without indication of its connection with the Planck law. Moreover, v appears in the normalised neutron mass $n \approx 1838.6836089(17)$ (13 ppb):

$$n^{1/3} \approx v\,(\pi/2)^2 \,. \tag{56}$$

The small deviation is attributed to a rationalisation of π involving again the heavy leaptons: 3, 7, 16, $-(1+\tau/\mu)^2$.

Another central constant in the Planck law is the irrational Apery constant $\xi(3) \approx 1.20205691$. The number of photons in a sphere of radius r is: $n_{ph}(r) = (4\pi/3)(r/l_{ph})^3$ with $l_{ph} = (hc/k_B\theta)(16\pi\xi(3))^{-1/3}$. The photon density is $l_{ph}^{-3} \approx 410.872$ photons/cm^3. The standard value is 410.7(4) cm^{-3} [17].

The critical photon/baryon ratio is $\eta_{cr} = n_{ph}(R)m_n/M$. While the number of photons exceeds the baryon number, it is the contrary for the energy densities, which is, for the CMB alone $u_{CMB} = (\pi^2/15)\hbar c/\lambda\!\!\!^{-4}_{CMB}$. However, the energy density of the sum CMB and CNB is the latter times $1 + 3 \times (7/8)(4/11)^{4/3} \approx 1.681321953$, to be compared to $u_{cr} = \rho_{cr}c^2$. One notes the dramatic relation between these two canonical ratios, with the 2 factor coming from photon polarisation (0.4 %):

$$\sqrt{2\eta_{cr}} \approx \frac{u_{cr}}{u_{CMB+CNB}} \,. \tag{57}$$

This is an Eddington-type relation, confirming that there are only three neutrinos, and ruining again the standard *evolutionary* cosmology. Moreover (0.08 %):

$$E = l_{ph}^{(CMB)}/\lambda\!\!\!^{-}_e \approx (\pi a^2)^2 \,. \tag{58}$$

This term is central in the unification number [29] (0.07 %):

$$U = \Phi^{137} \approx (1 - e^{-v})^{-1}\,(\pi a^2)^6 \,. \tag{59}$$

We recall that this quasi-whole number, based on the golden number Φ, shows a holic character [29] (0.03, 1, 0.07 %, 43 ppm, 0.4 %) :

$$U \approx (\pi P/Dp_K)^2 \approx E^3 \approx (pH/2a)^7 \approx (\tau^2/\mu^3)^{210}\,\sqrt{\delta} \tag{60}$$

with $D = 196883$ the Monster Moonshine dimension [43].

Atiyah introduced also the constant

$$\Gamma = \gamma a/\pi \tag{61}$$

as a simplification term. One observes:

$$2/\delta = 2a^3/pH \approx (1/2d_e)\ln(pH)/\ln a \approx \ln a/\ln\Gamma\,. \quad (62)$$

With $w = F/W$, this leads to (22 ppm): $a/\Gamma = \pi/\gamma \approx w^\delta$ while, with $z = F/Z$ (3 ppm): $137/\Gamma \approx z^{\sqrt{f}/2}$. Recall that $wz \approx \sqrt{a}$, while f is the Bizouard strong constant precising the inverse 8.44(5) of the standard "strong coupling constant" [17]:

$$f = a_w/2\pi(pH)^{3/2} \approx 8.43450\,. \quad (63)$$

In cosmology, Γ and the canonical e^π enter the following dramatic simplification of the above (Section 4.1) single-electron cosmic formula (0.3 ppm):

$$a' = ((\ln(R_1/\lambda\!\!\!^{-}_e)+\gamma-1)/(\pi^2/6-1)) \approx \ln(R/\lambda\!\!\!^{-}e)+\Gamma+e^\pi \quad (64)$$

so confirming the R value to 45 ppm.

Moreover, this confirms the role of $j = 8\pi^2/\ln 2$, the Sternheimer scale factor [3] (to 0.013 %, 0.013 %, 0.046 %):

$$j \approx \ln(R/\lambda\!\!\!^{-}e) + \Gamma \approx a - e^\pi \approx e^\pi \ln a\,. \quad (65)$$

The Titts group order $13 \times 2^{11}3^35^2$ [44] completes the biophysics relations involving central temperatures [3]:

$$j \approx T_{mam}/T_{CMB} \approx O_T/W \quad (66)$$

$$10^2 \approx T_{H_2O}/T_{CMB} \approx O_T/Z\,. \quad (67)$$

The pertinence of O_T is confirmed by the 2 ppb relation, where 71 is the biggest prime in the Monster order:

$$2 \times 137^2 + 21 = 23^2 \times 71 \approx 3 \times 137 d_e O_T/D\,. \quad (68)$$

The mammal wavelength enters (1%)

$$(Rl_P)^{1/2} \approx hc/kT_{mam}\,. \quad (69)$$

It is known that the reduced series $8k' + 2$ gives for $k' = 1$ and 3 the canonical values 10 and 26. Now the value $k' = 2$, $d = 18$ is at last interpreted: *the couple thermal photon-Life is at the upper center of the Topological Axis*, while the down center is the Higgs boson (Fig. 1). The real center, as seen above, is the dimension $d = 16$. Moreover, to 0.1%, the water triple point enters (0.1 and 1 %):

$$(R'l_P)^{1/2} \approx hc/k\theta_{H_2O} \quad (70)$$

$$\theta_{H_2} \times \theta_{O_2} \approx \theta_{H_2O} \times \theta_{CMB}\,. \quad (71)$$

This shows that chemistry is also involved [3].

The study of the 22 amino-acids [3] has shown that j is also a computation base. Indeed, to 2%: $j^{22} \approx 3P^2$ and, more precisely, to 0.01 % : $j^{22} \approx Pp_E^7$ where $p_E \approx 1847.599459$ is the Eddington mass ratio of the couple proton-electron, the roots ratio in the Eddington equation $10x^2 - 136x + 1 = 0$.

8.4 The ubiquity of a^a

Since 137 is a number of parameters, it must be interpreted as a dimension *i.e.* a privileged exponent. However, from the Computation Hypothesis, a must be an optimal base also. *So the term a^a must be central.*

Indeed, apart a π factor, a^a is the Grandcosmos volume with unit length the hydrogen radius, to 0.4 and 0.5 %:

$$(4\pi/3)(R_{GC}/r_H)^3 \approx a^a/\pi \approx 3(1/\ln 2)^{\sqrt{pH}}\,. \quad (72)$$

Note that the $\ln 2$ *factor involves information theory*. This relation is tied to the following property of the above unification factor (0.06 and 0.1 %):

$$U = \Phi^{137} \approx a^{p/a} \approx (1/\ln 2)^{pn/137^2}\,. \quad (73)$$

Moreover, the dramatic relation $a^a \approx e^{p/e}$ has been connected with the fifth optimal musical scale (306 notes) and to the operational definition of e [3]. Hence, we look here for its manifestations in classical mathematics.

The famous Lucas-Lehmer primality test uses the series of whole numbers $N_{n+1} = N_n^2 - 2$, starting from $N = 4 = u_3 + 1/u_3$, with $u_3 = \sqrt{3} + 2$. The latter is a special case of diophantine generators $u_n = \sqrt{n} + \sqrt{(n+1)}$, whose entire powers are close to whole numbers. One shows that $N_n \approx u_3^{(2^n)}$, and for n = 9:

$$u_3^{(2^9)} \approx (2(137^2 + 48))^{64} \approx a^a \quad (74)$$

defining a to 39 ppm and showing that the Rydberg term $2a^2$ plays a central role.

Also, with the Pell-Fermat generator $u_1 = 1 + \sqrt{2}$:

$$a^a \approx u_1^{3\times(2^8-1)} \quad (75)$$

which defines a to 0.3 ppm. So the number a establishes a connection between u_1 and u_3, two of the simplest arithmetic generators. This opens a *new research in pure mathematics*.

8.5 The intervention of sporadic groups

One observes, to 30 ppm, 0.5 % and 0.05 %:

$$O_M \approx (\ln\ln\ln O_M)^{2(136+d_e)} \approx (\pi/2)^{2a'd_e^2} \approx (F/af)^{20}\,. \quad (76)$$

Moreover (0.036 % and 0.038 %):

$$O_M^{1/10} \approx 495^2 \approx f(\gamma\Gamma) \quad (77)$$

where $495 = g_0/16$, implying the order g_0 of the smallest sporadic group (Mattieu) order M_{11}. Note that 495 is a unity less than the Green-Schwarz string dimension 496, the third perfect number, after 6 and 28. The precision 1.7 ppm of $f(\gamma\Gamma) \approx 495^2(a/137)$ suggests that the Higgs ratio is 495^2, corresponding to 125.175 GeV (Fig. 1 and Table 1).

The product of the 6 pariah group orders verifies (7 ppm):

$$\Pi_{pariah} \approx (F/a)^{20}/d_e^2 \tag{78}$$

thus, the above cosmic Tifft ratio F/a (Section 4.4) is directly tied to the six pariah groups. This establishes a connection between the six pariah groups and the Monster group (0.7 %):

$$\Pi_{pariah}/O_M \approx f^{20}\,. \tag{79}$$

These six pariah groups are not identified to form any family. By contrast, the 20 normal sporadic groups form the so-called *happy family* which is closely related to the Monster. The product of the 20 groups of the happy family shows, to 0.015%, 1% and 0.45 %:

$$\Pi_{happy} \approx \delta \times a^a \approx (j/495)^2\,\Gamma^{210} \tag{80}$$

where $j/495$ is close to the weak mixing angle 0.23116(12) [17], to 0.45 %. This confirms the above Complete Holic Principle, and the computation role of Γ. Moreover, to 2%: $a^a \approx \Gamma^{209}$. From the order of the Baby-Monster $O_B \approx \Gamma^{24}$, and $209 = 137 + 3 \times 24$ (1 and 2 %):

$$O_B \approx \Gamma^{24} \approx (a/\Gamma)^{a/3} \tag{81}$$

where $a/\Gamma = \pi/\gamma$ is the above canonical Atiyah ratio.

The total product of the 26 sporadic orders Π_{26} verifies (0.27 %):

$$\Pi_{26} \approx (9/2)(R_{GC}/\lambda\!\!\!^{-}_M)^2\,. \tag{82}$$

Now Π_{26} is close to the holic term $e^{4\times 210}$, whose a^{th} root is very remarkable (65ppm, 98 ppm, 5 ppb):

$$e^{4\times 210/a} \approx 2e^{2e} \approx H/4 \approx 26 \times (2 \times 26 + 1)/3\,. \tag{83}$$

Note that p/g_0 is close to the above weak mixing angle (0.3 %). This ratio appears as calculation base in the product of cardinal orders of the Monster and the baby-Monster groups, to 1%, 0.2 %, and 1 %:

$$O_M O_B \approx H^{2H/a} \approx (g_0/p)^a \approx (496/j)^{137} \tag{84}$$

confirming the central role of the weak mixing angle. The photon number in the visible universe is (0.1 % and 0.2 %):

$$n_{ph} \approx (3/\pi)\, e^{e^6/2} \approx \sqrt{\delta}\, O_M O_B\,. \tag{85}$$

With N_{ph} the photon number in the Grandcosmos, and $N_n = M_{GC}/m_n$ the equivalent neutron number in the Grandcosmos, one observes (3 %, 0.5 %):

$$\sqrt{N_{ph} N_n} \approx e^{n/3} \approx (O_M^3/U)^2 \tag{86}$$

confirming that the Grandcosmos is the external thermostat of the visible Universe. This is tied to (3 %, 0.08%, 2.5%,1%):

$$e^{137e} \approx U e^{n/6} \approx (e/3) e^{ea} \approx O_M^3 \approx 496^{60}\,. \tag{87}$$

With the tachyonic ratio $V = R_{GC}/R = C/c$, the orders of the two giant sporadic groups enter (0.2 %, 0.1 % and 79 ppm):

$$V \approx 44\pi N_S \approx (a/\pi) O_M D \approx (a/\pi) O_B P a^{3/2} \tag{88}$$

where $N_S = 2^{65} \times 3^{41} \times 5^{28}$ is the Systema number [45].

The corrected Eddington's number $N'_{Ed} = a \times 2^{256}$, where 136 is replaced by a, shows (4.5 ppm and 0.03 %):

$$N'_{Ed} \approx 6 \times 137 P O_M \approx (3/4) a p a_w (V/O_M)^9\,. \tag{89}$$

With the 4D area $s_4 = 2\pi^2 a^3$, the holic reduction

$$(R/\lambda\!\!\!^{-}_e)^7 \approx (3/2) O_M^5 \approx s_4^{35} \tag{90}$$

implies $O_M^{1/7} \approx s_4$. Indeed, the Monster appears to be close to the seventh power of the pariah group J_3 (0.2 ppm):

$$O_M \approx d_e J_3^7 \sqrt{p/p_0}\,. \tag{91}$$

The above relations proves that physics establish unexpected bridges between sporadic groups, including the Titts one.

9 The fine-tuning with basic mathematical constants

We look here for relations involving basic mathematical constants, noting firstly that, to 6.5 ppm: $p \approx \Gamma(\pi e)^2$.

9.1 The optimal calculation base e confirmed

The electron magnetic moment $2d_e$ appears in (0.7 ppm):

$$a/\Gamma = \pi/\gamma \approx 2d_e \times e\,(p_0/p)^2\,. \tag{92}$$

The Topological Axis shows clearly that the Grandcosmos is defined by the following conjunction (1%):

$$f(k = e^2) = \exp(2^{e^{2+1/2}}) \approx \exp(e^{2e} + e^2) \tag{93}$$

where the supplementary term $\exp(e^2)$ is close to $a^{3/2}$. Note the following properties of the "economic number" e^{e^e}, to 0.4 %, 6 ppm and 0.8 ppm:

$$e^{e^e} \approx (\ln p)^{\ln p} \approx 137(e^e)^3 \approx e^e a \sqrt{pH}(p/p_0)^2\,. \tag{94}$$

With $a_1 = a - 1$ (8 ppm, 0.2 ppm, and 0.05 %):

$$e^{e^e}/a_1^2 \approx 4 \ln P \approx a \ln(9/2) \approx 5^{2^7} \tag{95}$$

showing the role of musical bases 2, 3 and 5. Note that the Topological Axis terminal term e^2 is the limit of the following musical series:

$$(3/2)^5 \approx (4/3)^7 \approx (5/4)^9 \approx (6/5)^{11} \approx ... \approx (1 + 1/n)^{2n+1}$$

a series converging more rapidly than the classical $(1 + 1/n)^n$. The first two terms defines the occidental 12 tones scale. Note that, to 0.6 % and 0.03 %:

$$R/\lambda\!\!\!^{-}_e \approx 2^{2^7} \tag{96}$$

$$R'/\lambda\!\!\!^{-}_e \approx (3^3)^{(3^3)} . \quad (97)$$

The canonical ratio $R_{GC}/\lambda\!\!\!^{-}_M = 2P^9/a^6 pH$ confirms the Full Holographic Principle, to 0.04 %:

$$R_{GC}/\lambda\!\!\!^{-}_M \approx (137e/a)^{2\times 210} \quad (98)$$

exhibiting (0.3 ppm): $(a/137)^{420} \approx (137-3)/120$ with $137-3 = 7+127$ showing the Combinatorial Hierarchy terms [3].

9.2 The Lenz-Wyler formula

Wyler published a value approaching a to 0.6 ppm and confirmed the pertinence of the Lenz approximation which plays a central role above: $p_0 = 6\pi^5 \approx p$ to 18.824 ppm.

The Lenz-Wyler formula is the product of the area by the volume of a 3D cube with side π. If one considers a 3D cube with side 5, privileging again the identification dimension = exponent, this gives $6\times 5^5 = 137^2 - 19$. This is not a chance coincidence because this relation has long time been deduced from basic considerations on quarks [29]. Indeed with $u = 5$ and $d = 6$, the combination $uud = 150$, whose power 3/2 is close to H, while the combination $udd \approx (n/a)^2$ shows the neutron/electron mass ratio n. This leads to (0.012 %) $6\times 5^5 \approx (aH/n)^2$. Note that, with $q = 2^{12}$ to 0.03 %, 2.5 % and 41 ppm:

$$R_{GC}/\lambda\!\!\!^{-}_e \approx q\times 5^{137} \approx 6^{137}/q^2 \approx 6^{128}/(1+1/\sqrt{2}) .$$

Since $R/\lambda\!\!\!^{-}_e \approx 2^{128}$, *the factorisation of 6 leads to a natural Universe-Grandcosmos partition*, and to the following approximation for the tachyonic celerity ratio (0.01%)

$$U = C/c \approx 3^{128}(p_K/p_G\delta)^2$$

where p_K is the Koide-Sanchez constant (see Section 9.5). This confirms the role of the correspondence quark up = 5 and *quark d = 6 with a double structure*. This elimination of q leads to (2.6 %): $(R_{GC}/\lambda\!\!\!^{-}_e)^3 \approx (uud)^{137}$.

It is an example of *immergence*, *i.e.* deducing the small from the large, in a striking similitude between cosmology and nuclear physics. Another example was encountered in Section 2.4, where dimensional analysis gives the visible Universe radius, in an easier way than the equivalent one for the hydrogen atom radius, since for this case there is no evidence that c must be left out. Another example signals a general misconception: the coherence of the stimulated emission in a laser is a *global effect* in a homogeneous media (atomic coherence).

9.3 The Archimedes constant π as a calculation base

From (27), the value of the topological function for the main string dimension 26 renders, to 0.1%, the same Lenz-Wyler form $f(26) \approx 6(2\pi^2a^3)^5$, where $2\pi^2a^3$ is the area of a 4-sphere of radius a. Moreover, with n/p the mass ratio neutron/proton, to 0/3%, 0.02% and 1 ppm:

$$(p/n)(R/\lambda\!\!\!^{-}_e)^2 \approx (f(26)/6)^2 \approx (2\pi^2a^3)^{10} \approx \pi^{155} . \quad (99)$$

The corresponding approximation π_R of π shows the fractional series 3, 7, 16, $-u$, with $u \approx 2\times 137$, confirming again the rationalization hypothesis of Section 3. This leads to the rational value $\pi_R = (355u-22)/(113u-7)$. This corresponds to the above G value to 10 ppb accuracy. Since $(R/\lambda\!\!\!^{-}_e)^2 \approx 2^{256}$, this illustrates the following musical relation involving again 137: $2^{1/155} \approx \pi^{1/256} \approx (2\pi)^{1/3\times 137}$. The scale with 155 notes is not known, but 137 appears also in the classical musical scales [3]. Whole powers of π appear in the even order Riemann series, and in: $a \approx 4\pi^3 + \pi^2 + \pi$ (Reilly formula, 2 ppm), while $a \approx \pi^{9/2}2^{-1/3}$ (8 ppm). Moreover, with $P = \lambda\!\!\!^{-}_e/l_P$ (0.3 and 0.07 %):

$$P^3 \approx \pi^{a-2} \approx (2\pi R/\lambda\!\!\!^{-}_e)(2\pi l_K/r_e) \quad (100)$$

confirming the Planck volume and the Kotov length.

9.4 The four forces connection in ppb fine-tuning

The particle standard model achieved the unification between electromagnetism and weak nuclear force, with extension to strong nuclear force in the Grand Unification Theory (GUT), but without any synthesis with gravitational force. However, the Topological Axis shows clearly that GUT gauge boson with 2.3×10^{16} GeV seems confirmed. Very precisely, in Section 4.2, it is proven that the CCO oscillation reveals a symmetry between the electroweak and gravitational forces. So we look here for a precise relation involving the 4 force parameters, a (electric), a_w (weak nuclear), f (strong nuclear) and a_G (gravitation). The later force is equivalently represented by $p_G = P/2^{127/2}$, with $P = m_P/m_e$.

With the Atiyah constant $\Gamma = \gamma a/\pi$ (Section 8.2), inside the 0.5 ppm measurement precision: $a_w = F^2 = (137\times 2\Gamma)^3$. Now a_w is a cube: $a_w = (\lambda\!\!\!^{-}_e/l_{eF})^3$, with $l_{eF} = (G_F/m_ec^2)^{1/3}$:

$$\lambda\!\!\!^{-}_e/l_{eF} \approx 137\times 2\Gamma \quad (101)$$

$$F = a_w^{1/2} = E_F/m_ec^2 \approx 573007.3652 \quad (102)$$

$$aF/\sqrt{(pH)} = 2\pi a f pH/F \approx \pi(4n/\Gamma)^3/p_G \approx \mu^2 \quad (103)$$

where μ is the muon/electron mass ratio, *inside its 20 ppb undetermination*, so proposing the value:

$$\mu \approx 206.7682869 . \quad (104)$$

Note that $4n/\Gamma$ is close (3.4 ppm) to the monstrous 5th term 292.6345909 in the fractional development of π which is itself very close to $n/2\pi$ to 3.4 ppm. Since the fractional development of π is to this date an unsolved problem, *this confirms that current mathematics is incomplete and that Nature uses rational approximations of* π. From the Koide relation, the corresponding value is $\tau \approx 3477.441701$, tied to the economic number (0.6 ppm):

$$e^{e^e} \approx \tau(2a)^3/137^2 . \quad (105)$$

From $\tau \approx e^{3e}$ and $8a \approx e^7$, this illustrates the reduction $e^e \approx 7 + 3e$. The pertinence of the economic number is confirmed.

The corresponding Koide-Sanchez constant is

$$p_K \approx 1842.604994\,.$$

This leads to three ppb relations, where $\pi_a = (355u + 22)/(113u + 7)$, with $u = a\sqrt{2}/3$, and $H_e = 8e^{2e}$ is the economic 33 ppm approximation of H:

$$p_K^4/pH \approx (4\pi_a)^4 a \approx (p_G H_e/aH)^4 D^2/n(D+1) \qquad (106)$$

$$n\tau/2 \approx HH_e(D/(D+1))^3(p_K/p_G)^9 \qquad (107)$$

where D and $D + 1$ are the characteristic numbers of the Moonshine correlation [43]. This confirms the Eddington symmetry hydrogen-tau lepton [5].

The above relations show a dual form, the first one without any numerical factor:

$$ap_G/\pi\sqrt{(pH)} \approx (nF/137^2\Gamma^3)^3 \approx (4n/\Gamma)^3/F\,. \qquad (108)$$

Now, as was recalled above, the exponents represent the number of dimensions. So, this represents a dimensional reduction, eliminating 137, from 9-D and 6-D to 3-D, which could be associated with the superstring theory, where the equations are coherent only if space has 9 dimensions, and if the 6 supplementary dimensions unfold on very small distances [47].

The following weak boson ratios W and Z match (1): $R/\sqrt{(\lambda\!\!\!^{-}_p \lambda\!\!\!^{-}_H)} \approx (WZ)^4$ *in the ppb range*:

$$W \approx 137^2\Gamma/3d_e \qquad (109)$$

$$Z \approx ap^2\pi^4/137d_e n\,. \qquad (110)$$

The ultimate theory must explain these ppb relations.

10 Discussion

For many, cosmology is the hardest chapter of physics. This modern negative opinion is in fact in contrast with the ancient culture, for which *the cosmology is the first of all sciences, so must be the simplest*. In the original meaning of the word "revolution", this article is a return to the source of science, the "all is whole number" of Pythagoras. Even the degenerate form of topological or holographic relations, the simplest diophantine equations, the Holic Principle, shows direct pertinence. In particular, it emphasizes the 30 dimensions, which appear decisive in the Topological Axis, and are identified with the sum of 26 string dimensions and 4 of usual spacetime.

The distinction between length and time must be emphasized, as Poincaré, the father of 4-D relativity theory recommended [25]. Indeed their confusion, by writing $c = 1$, impeded the fact that the Hubble-Lemaître radius R is a trivial length, directly given by the prospective c-free dimensional analysis, which gives also the cosmic temperature (37) and the cosmic supercycle period (22).

This means also that the International System must go back to only three fundamental unities, Mass, Length and Time.

The Hierarchy and Computation principles presented in Section 1 are confirmed both by the Topological axis, the geodimensional Universe-Grandcosmos couple, and the monomial relations (*i.e.* merely products of parameters). These accurate monomial relations reunify mathematics and physics. The precision reaches the ppb domain: they cannot be due to chance. This shows how the so-called "free parameters" are misnamed: they are imposed by Nature proving the Cosmos unicity. As Atiyah wrote, rather misleadingly [42]:

> Nobody has ever wondered what the Universe would be if π were not equal to 3.14159.... Similarly no one should be worried what the Universe would be if a were not 137.035999...

In fact a must be rational, and the mathematical π is illusion. Nevertheless, this article is a *definite refutation* of the Multiverse hypothesis. In this respect, the high precision in the measurement of the electric and Fermi constants, proton, neutron and muon masses, Kotov cosmic period, and, with lesser precision, the background temperature, must be saluted as decisive achievements.

The pertinence of these simple monomial relations cannot be admitted by the standard community, arguing for instance that since the proton is composite, its mass cannot enter simple relations. The same argument is presented for the theoretical dependence of the electric constant a with other constants, or with the energy level. These are reductionist arguments, unable to explain the fine-tuning phenomena, and leading to the sterile concept of unexplained emergences. By contrast, the holistic approach implies the concept of *immergence*, resulting from the ancestral idea that Cosmos simplicity is the real origin of science. It is strange, revealing and troubling that this term *immergence* is a neologism.

The Cosmos concept has long been forgotten. This is the reason why quantum physics is not really understood. Indeed, the simple fact that the propagation of anything, light or matter, is wavy, while the reception is a quantum, was a central mystery along the last century. This simple fact induces non-locality, so the necessary intervention of cosmology. Moreover, the optimal utilisation of the wavy propagation is holography, whose formalism is similar to the quantum one. Thus it is logical to find holographic relations in cosmology. Moreover, the similitude between the formalisms of quantum physics and holography is so tight that the double-step holography is similar to the double step of any interaction: tachyonic propagation – non-local cosmic optimisation – local quantum reception.

Thus tachyonic-holography physics is necessary. Hence, it was an error to reject the bosonic string theory under the pretext it involves tachyons [49]. Quite the contrary, it is an essential advantage. This is confirmed by the central impor-

tance of the bosonic dimension $d = 26$ in the Topological Axis, which is nothing that the extension to smaller numbers of the Double Large Number coincidence, that only Eddington interpreted correctly, by rejecting the single Bang model. Many invoked the temporal variation of the parameters, which is a negation of the idea that physics have universal laws. Finally, the expedient of the Anthropic Principle was imposed to the community by some leaders: this is definitely refuted in this article.

Moreover, the standard Holographic Principle must be generalized to wavelengths other than the Planck length, in particular the topon, the visible Universe wavelength, in 1-D holography, which breaks by an enormous factor, about 10^{61} a taboo of current thinking: the Planck wall, resolving the vacuum energy dilemma factor 10^{122}, and sustaining the Oscillatory bounce model which unifies the two main cosmologies.

This leads back to the main hypothesis of this article: the Cosmos is a computer, and the dimensionless parameters are calculation bases. A common point with the brain is precisely this *multibase* character, experienced in musical sensation. It is no chance that the parameters are encountered in the musical scales and DNA chain. Thus, intelligent life receives a justification: to help the cosmological computation. This Inverted Anthropic Principle answers the first of all questions: *why one asks questions?*

Thus, *intelligent life must be universal*. The famous Fermi question "where are they?" is not a paradox, since any abnormal observation is *a priori* rejected by a dogmatic community. This destroys the Darwin "accidental life" approach, a generally admitted so-called "theory" with too much missing links [48].

The same rejection seems to apply now to the Sternheimer "scale wave" and Atiyah's last work. The present article shows that at least parts of these works are very pertinent. This follows the rejection (with the notable exception of Schrödinger) of Eddington [5] himself. Only Eddington interpreted rightly the Cosmic Large Number correlations, as recalled in this article. While he dared to apply the exclusion principle in cosmology, it is the basis of our single electron cosmologic model (Section 4.1) which rehabilitates once more his work. Also, fortunately, the large theoretical advance of Eddington is now recognized [51], but without mentioning a crucial point: *he predicted the tau fermion with a right order of mass*, 30 years before its surprising discovery, calling it heavy mesotron [5]. Moreover, it seems that no one realizes that the Eddington prediction for the baryon number in the visible Universe is so accurate. Note that many mocked the Eddington Large Number, not to speak of his number 137, completely rehabilitated by the monomial relations.

However, curiously, Eddington believed in the Copenhagen statistical interpretation. Thus, he did not reach the above conclusions. At his epoch the holography was not yet discovered: *it is a strange, and revealing, fact of science history that this essential property of wave propagation was so lately discovered* [33]. However, with his Large Number which fits so well the cosmic neutron population, Eddington anticipated the present physics-arithmetic fusion and its touchstone, the Holic Principle.

11 Conclusions: cosmic simplicity at work

The present article confirms the Topological Axis, which was obtained by the simplest visualizing method to represent in a single figure the characteristic lengths in macro and microphysics, taking the electron reduced Compton wavelength as unity. *The double logarithm representation was the simplest one, and it appeared later that this was the reunion of a series of height 1D-2D holographic relations, respecting the topologico-algebraic Bott sequence*.

The application of the old direct scientific method, looking for fine tuning between physical parameters leads to a return to the Perfect Cosmological Principle implying a steady-state Cosmos, confirmed by holographic relations. The standard cosmological principle was unduly limited to spatial homogeneity. The relativity theory, unable to define an inertial frame, is a local one and do not apply to cosmology at large: the absolute space is reestablished, realized by the Microwave Cosmic Background, which identifies with the Grandcosmos frame. Meanwhile, the Kotov period is an absolute clock, the *déphasage* of coherent oscillations between quasars being ruled by the tachyonic celerity.

The simplest model, the gravitational hydrogen molecule gives the Hubble radius R, explaining the 2 factor and justifying the elimination of c, as in the hydrogen atom Haas-Bohr model [3]. This corresponds to a Hubble constant 70.790 (km/s)/Megaparsec, consistent with the recent measurement [6]: 72(3) Megaparsec/(km/s), which confirms the direct novae measurement, but disagrees (3σ) with the standard value.

The simplest statistical theory of Eddington gave another justification to R. Also, particularly simple and elegant is the Large Eddington number, giving correctly the number of neutrons in the trivial fraction $3M/10$ of the observable universe, *probably the most dramatic prediction in scientific history*.

The simplest proof of the computation basis character of the electrical parameter a is provided by the multiple appearance of the terms e^a and a^a.

The profound significance of a number of dimensions is the number of independent variables, which is a fundamental invariant, whatever the theory [54]. So, it is logical to advance a hypothesis that 26 physical parameters are defined by the 26 sporadic cardinal orders. Since Sporadic Groups are associated with octonion algebra [55], this rejoins a prediction of Atiyah's last work, the essential role of octonion algebra in the final theory [42].

The problem of the stability of the solar system must be revisited, taking into account seriously a cosmic influence, characterized by the Kotov period and length. Also the Pi-

oneer, Tifft and Arp effects must be seriously considered, guided by the flickering time-length-mass concept.

This article answers several main problems:

- 1/ Unification gravitation-quantum physics, by rehabilitating the forgotten Eddington statistical theory.
- 2/ The real significance of quantum physics, by assuming physics is based on arithmetics.
- 3/ The overall unification by showing that cosmology is the basis of united science.
- 4/ The role of dimensionless parameters, by proving that they are optimal basis of computation tied with the Holographic Principle and its arithmetic form, the Holic Principle, which explains why normal space has 3 dimensions.
- 5/ The necessity of the Cosmos vasteness resulting from holographic scanning and the rationalization of e and π.
- 6/ The acceleration of expansion, which was predicted by the Eddington *invariant* cosmological constant $1/R^2$, is tied to a repulsive force proportional to distance, leading to *exponential* recession. There is no need of the so-called "dark energy".
- 7/ The very existence of dark matter is proven, from the number of neutrons in the trivial fraction 3/10 of the visible Universe critical mass, which identifies with the very symmetric Eddington number 136×2^{256}. *The nature of dark matter would be simply a matter-antimatter oscillation in phase quadrature with the ordinary one* [3].
- 8/ The introduction of the topon in the Holographic Principle justifies at last the 10^{122} gap between vacuum energy and that of the visible Universe.
- 9/ The Grandcosmos is huge, but not infinite, in conformity with the Cosmological Computational Principle.

In short, the rediscovered Cosmos unifies the two main modern cosmologies in a rapid matter- antimatter oscillatory bounce. The Cosmos appears as *simple, unique, permanent, computational, deterministic, transplanckian, cyclic, topological and inverse-anthropic*. It is now clear that present mathematics and particle physics are incomplete, and this Coherent Cosmology announces a reunification of *philosophy, mathematics, physics, chemistry, computational science and biology*. In particular, the pre-Socratic Parmenide philosophy of permanence must be reconsidered favorably.

12 Predictions

This article leads to many predictions, in particular:

- 1/ The very large infrared telescopes will show in the very far field old galaxies instead of expected young ones. Then no artifice, such as inflation, dark energy, multiverse, ..., will not save the standard evolutionary model, based on the imperfect cosmological principle.
- 2/ The CMB temperature and the baryon mean density will appear temporal invariant.
- 3/ The particle physics will integrate the Koide relation together with the Koide-Sanchez constant, and introduce composite quark down and massive photon, graviton, gluons and string. Also the supersymmetry will restablish the Eddington connection proton-tau.
- 4/ The computational software should be boosted by the principle of multibase computation.
- 5/ The DNA chain will reveal as a 1-D temporal hologram, see [52].
- 6/ The Lucas-Lehmer series, in connection with the canonical generators ($\sqrt{n} + \sqrt{(n+1)}$), especially the Planck-Fermat one ($1 + \sqrt{2}$) will define a.
- 7/ The 26 sporadic groups as well as the Titts one will reveal determinant in the Ultimate Theory.
- 8/ The Eddington Fundamental Theory will be revisited, especially the genesis of his Large Number, so clearly tied to the 16×16 symmetric matrix.
- 9/ The Combinatorial Hierarchy [41] and Moulin systemic approach [45] will be reconsidered.

Received on April 2, 2019

References

1. Carr B. J. and Rees M. J. The anthropic principle and the structure of the physical world. *Nature*, 1979, v. 278, 605–612.
2. Borcherds R. Monstruous Moonshine and Monstruous Lie Superalgebras. *Invent. Math.*, 1992, v. 109, 405–444.
3. Sanchez F. M. A Coherent Resonant Cosmology Approach and Its Implications in Microphysics and Biophysics. *Progress in Theoretical Chemistry and Physics*, 2017, v. 30, 375–407, DOI 10.1007/978-3-319-50255-7-23. Sanchez F. M. Coherent Cosmology. vixra: 1601.0011. Sanchez F. M., Kotov V. and Bizouard C. Towards Coherent Cosmology. *Galilean Electrodynamics*, 2013, special issue, 63–80.
4. Poincaré H. Sur la théorie des quanta. *J. de Physique*, 1912, v. 2, 5.
5. Eddington A. S. The Fundamental Theory. Cambridge, 1946.
6. Bonvin V. *et al*. H0LiCOW-V. New COSMOGRAIL time delays of HE0435-1223: H0 to 3.8% precision from strong lensing in a flat ΛCDM model. *Monthly Notices of the Royal Astronomical Society*, 2017, v. 465 (4), 4914–4930. arXiv: astro-ph/1607.01790v2.
7. Bondi H. and Gold T. The steady-state theory of the expanding universe. *Monthly Notices of the Royal Astronomical Society*, 1948, v. 108, 252.
8. Hoyle F. A new model for the expanding universe. *Monthly Notices of the Royal Astronomical Society*, 1948, v. 108, 272–382.
9. Poincaré H. Dernières pensées. Flammarion, 1913, pp 102–103.
10. Sanchez F. M., Kotov V. A. and Bizouard C. Towards a synthesis of two cosmologies: the steady-state flickering Universe. *J. Cosmology*, 2011, v. 17, 7225–7237.
11. Davie P. The Accidental Universe. *C.U.P.*, 1993, 92.
12. Bousso R.The Holographic Principle. *Reviews of Modern Physics*, 2002, v. 74, 834.

13. Duplantier B. Introduction a l'effet Casimir. *Séminaire Poincaré*, 2002, v. 1, 41–54.
14. Damour T. The Entropy of Black Hole. *Sem. Poincaré*, 2003, v. 2, 89–115.
15. Nambu H. An empirical Mass Spectrum of Elementary Particles. *Prog. Theor. Phys.*, 1952, v. 7 (5), 595–6.
16. Kotov V. A. and Lyuty V. M. The 160-min. periodicity in the optical and X-ray observations of extragalactic objects. *Compt. Rend. Acad. Sci. Paris*, 1990, v. 310, Ser. II, 743–748. Fossat E., Boumier P., Corbard T., et al. Asymptotic g modes: Evidence for a rapid rotation of the solar core. *Astron. Astrophys.*, 2017, v. 604, A40, 1–17. DOI: 10.1051/0004-6361/201730460. Grec G., Fossat E. Calculation of pseudo solar narrow band oscillations produced by atmospheric differential extinction. *Astron. Astrophys.*, 1979, v. 77, 351–353. Kotov V. A. Evolution of the Sun and the Earth: the (un)known period 1.035 years. *Izv. Krym. Astrofiz. Obs.*, 2013, v. 109 (1), 232–253. Kotov V. A. Fast spinning of planets. *Earth Moon Planets*, 2018, v. 122 (1), 43–52. DOI:10.1007/s11038-018-9520-6. Sevin É. Sur la structure du système solaire. *Compt. Rend. Acad. Sci. Paris*, 1946, v. 222, 220–221. Kotov V. A. Motion of the fast exoplanets. *Astrophys. Space Sci.*, 2018, v. 363 (3), 1–5. DOI: 10.1007/s10509-018-3278-1.
17. Tanabashi M. *et al*. (Particle Data Group). The review of particle physics. *Phys. Rev. D*, 2018, v. 98, 030001. http://pdg.lbl.gov.
18. Quinn T., Speake C., Parks H., Davis R. The BIPM measurements of the Newtonian constant of gravitation, G. *Phil.Trans. R. Soc. A*, 2014, v. 372, 20140032. http://dx.doi.org/10.1098/rsta.2014.0032
19. Aschbacher M. Sporadic Groups. Cambridge Univ. Press, 1994.
20. Sanchez F. M. Towards the grand unified Holic Theory. In Pecker J.-C. and Narlikar J., eds. Current Issues in Cosmology. Cambridge Univ. Press, 2006, p. 257–260.
21. Kotov V. A. and Sanchez F. M. Solar 22 years cycle. *Astrophys. Space Sci.*, 2017, v. 362 (6), 1–6. DOI: 10.1007/s10509-016-2985-8.
22. Tifft, W. G. Redshift periodicities, The Galaxy-Quasar Connection. *Astrophysics and Space Science*, 2006, v. 285 (2), 429.
23. Arp H. The origin of Companion Galaxies. *Astrophysical Journal*, 1998, v. 496, 661–669.
24. Nieto M. and Anderson J. Using Early Data to Illuminate the Pioneer Anomaly. *Class. Quant. Grav.*, 2005, v. 22, 5345–5354. arXiv: gr-qc/0507052.
25. Poincaré H. La mécanique nouvelle. Gauthiers-Villars, 1924. Jacques Gabay, 1989.
26. Bars I. Gauge Duality, Conformal Symmetry, and Space-Time with Two Times. *Phys. Rev. D*, 1998, v. 58. arXiv: hep-th/9803188.
27. Cartan H. Démonstration homologique des théorèmes de périodicité de Bott I. *Séminaire Henri Cartan*, 1959-1960, v. 12 (2), n° 16, 1–16.
28. Grosmann M. and Meyrueis P. Optics and Photonics Applied to Communication and Processing. *SPIE*, Jan. 1979.
29. Sanchez F M. Holic Principle: The Coherence of the Universe. Entelechies, 16th ANPA, Sept. 1995, p, 324–344.
30. Sanchez F. M. Cohérence Temporelle d'ordre supérieur d'un laser déclench multimode. Doctorate Thesis, Orsay, number 1478, 1975, p. 11.
31. Okun L. B. Photon: history, mass, charge. Int. Conf. on the Photon, Warsaw University, 2005. arXiv: hep-ph/0602036.
32. Einstein A., Podolski B. and Rosen N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete. *Phys. Rev.*, 1935, v. 47, 777.
33. Gabor D. A new microscopic principle. *Nature*, 1948, v. 161, 777–778.
34. Marchal C. Physics with photons of non-zero rest mass. Proceedings of 28th Intern. Workshop, Protvino, Russia, 2005, p. 152–166.
35. Preskill J. Do Black Holes Destroy Information? International Symposium on Black Holes, Membranes, Wormholes, and Superstrings. arXiv: hep-th/9209058.
36. Nikolic H. Resolving the black-hole information paradox by treating time on an equal footing with space. *Physics Letters B*, 2009, v. 678 (2), 218–221. arXiv: gr-qc/0905.0538.
37. Carlson C. E. The proton radius puzzle. *Progress in Particle and Nuclear Physics*, 2015, v. 82, 59–77.
38. Witten E. The Three-Dimensional Gravity Revisited. arXiv: hep-th/0706.3359.
39. Sanchez F. M. Electrical Moonshine. vixra: 1801.0067.
40. Koide Y. Fermion-Boson Two-Body Model of Quarks and Leptons and Cabibbo Mixing. *Lett. Nuovo Cimento*, 1982, v. 34, 201.
41. Bastin T. and Kilmister C. W Combinatorial Physics. World Scientific, 1995.
42. Atiyah M. The fine-structure constant, 4th Heidelberg Laureate Forum conference (2018). https://www.heidelberg-laureate-forum.org/blog/video/lecture-monday-september-24-2018-sir-michael-francis-atiyah/.
43. Conway J. H. and Norton S. P. Monstrous Moonshine. *Bull. London Math. Soc.*, 1979, v. 11 (3), 308–339.
44. Titts J. Algebraic and abstract simple groups. *Annals of Mathematics, second series*, 1964, v. 80, 313–329.
45. Moulin T. Utilisation en physique et biologie de référentiels spatio-structuro-temporels engendrés par des relateurs arithmétiques. 11eme Congrès International de Cybernétique, Symposium 4, Namur, (1986).
46. Brannen C. A. The Lepton Masses. http://brannenworks.com /MASSES2.pdf, 2006.
47. Polchinski J. String Theory, Vol 1. Cambridge University Press, 1998, p. 22.
48. Chauvin R. Le Darwinisme ou la fin d'un mythe. ed. du Rocher, 1997
49. Woit P. Not Even Wrong: The Failure of String Theory and the Search for Unity in Physical Law. Basic books, 2006
50. Larin S. A. Quantum Chromodynamics with massive gluons. arXiv: hep-ph/1304.8107.
51. Durham I. T. Sir Arthur Eddington and the Foundations of Modern Physics. 2006, p. 111. arXiv: quant-ph/0603146v1 .
52. Widom A., Swain J., Srivastava Y. N., S. Sivasubramanian S. Electromagnetic signals from bacterial DNA. arXiv: physics.gen-ph/1104.3113v2.
53. Salingaros N. Some remarks on the algebra of Eddington's E Numbers. *Foundations of Physics*, 1985, v. 15 (6), 683—691.
54. Weigel D., Veysseyre R. and Carel C. Sur les symboles du groupe d'espace d'une wüstite de tri- incommensurabilité cubique et sur les groupes de Bravais de sa famille cristalline dans l'espace euclidien à six dimensions. *C.R. Acad. Sci. Paris*, 1987, v. 305, ser. II, 349–352.
55. Atiyah M. private communication, December 2018.

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