Issue 3 2016 | Volume 12

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

April–July 2016 **Vol. 12, Issue 3**

CONTENTS

Yurkin A.V. On the Descriptive Geometric Interpretation of Pauli Principle, Elements of the Mendeleev Table of Chemical Elements, and the Newtonian Laminar Current of a Liquid 149

Eid S.A., Diab S.M. X(5) Symmetry to ^{152}Sm 170

Johnson R.J. A Re-examination of Kirchhoff's Law of Thermal Radiation in Relation to Recent Criticisms (*Letters to Progress in Physics*) 175

Robitaille P.-M. A Re-examination of Kirchhoff's Law of Thermal Radiation in Relation to Recent Criticisms: Reply (*Letters to Progress in Physics*) 184

Rybicki M. Errata to "Mansouri-Sexl Test Theory: The Question of Equivalence between Special Relativity and Ether Theories" (*Letters to Progress in Physics*) 204

Esmail S.H. and Taha M.M. Application of the Differential Transform Method to the Advection-Diffusion Equation in Three-Dimensions 205

Millette P.A. On the Applicability of Bell's Inequality 211

Spivey R.J. Coincident Down-chirps in GW150914 Betray the Absence of Event Horizons 216

Marshall T.W. Repulsive Gravity in the Oppenheimer-Snyder Collapsar 219

Daywitt W.C. The Dirac-Electron Vacuum Wave 222

Muralidhar K. Mass of a Charged Particle with Complex Structure in Zeropoint Field . . 224

Brodet E. The Relationship Between the Possibility of a Hidden Variable in Time and the Uncertainty Principle 231

Daywitt W.C. A Modern Interpretation of the Dirac-Electron Continuity Equation 234

Marshall T.W. Optics of the Event Horizon Telescope 236

Matwi M. The Dual Behavior of Quantum Fields and the Big Bang 241

Consiglio J. On Quantization and the Resonance Paths 259

Belyakov A.V. On the Nature of Ball Lightning 276

Feinstein C.A. Dialogue Concerning the Two Chief World Views (*Letters to Progress in Physics*) 280

Spivey R.J. Criteria for Aerial Locomotion in Exoplanetary Atmospheres: Revisiting the Habitable Zone for Flying Lifeforms 284

Muralidhar K. The Structure of the Photon in Complex Vector Space 291

Obagboye L.F., Howusu S.X.K., Chifu E.N. and Omaghali E.J.N. Gravitational Waves from a Sinusoidially Varying Spherical Distribution of Mass 297

Daywitt W.C. Gravitational Shielding as Viewed in the Planck Vacuum Theory 301

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*.)

On the Descriptive Geometric Interpretation of Pauli Principle, Elements of the Mendeleev Table of Chemical Elements, and the Newtonian Laminar Current of a Liquid

Alexander V. Yurkin
Puschino, Russia. E-mail: alvl1yurkin@rambler.ru

This work presents a two-dimensional and three-dimensional geometrical research of a ray system. We consider trajectories of motion of the particles having a half-integer spin. Interpretation of Pauli Principle showing distribution of electrons on power levels of the atom is given herein. The number of the electron shells in our model of the atom doesn't exceed 8. We give a geometric interpretation of the main, azimuthally, magnetic and spin numbers in the form of angles and distances. We show forth that the hyperbolic dependence of energy on the main quantum number n of the hydrogen atom ($E_n \sim -1/n^2$) known from experimental spectral studies, Bohr's theory and Quantum Mechanics can also be obtained from our geometrical formulation of Pauli Principle. Also, in the framework of research of the suggested ray model, the step structure of the layers at a laminar current of a liquid is deduced.

Contents

Introduction 149
1. Half-integer system of eight groups of rays 150
1.1. Two-dimensional projection of Gaussian (paraxial) rays 150
1.2. Three-dimensional projection of Gaussian (paraxial) rays 150
1.3. Periodic and acyclic trajectories 150
1.4. Periodic trajectories and step layers in the laminar current of a liquid 152
2. Gaussian (paraxial) rays and Pauli Principle 159
2.1. Angles, distances and quantum system 159
2.2. Periodic tables and geometrical constructions 160
2.2.1. Creation of the first shell of a quantum system 160
2.2.2. Creation of the second shell of a quantum system 160
2.2.3. Creation of the third and fourth shells of a quantum system 160
2.2.4. Pauli Principle and the geometric system of the hydrogen atom 162
Conclusions 167
Acknowledgements 167
References 167
Appendix 168

Introduction

Descriptive geometric models are used for the evident description of various phenomena, including quantum phenomena [1].

In works [2, 3], we already introduced a geometric model on the plane consisting of systems of paraxial rays describing distribution of light in lasers, turbulent and laminar flows of a liquid on pipes, and also finding an electron in the infinite deep potential. In work [3] we noted that the aforementioned model can be used for a descriptive interpretation of moving particles with the integer or half-integer spin.

In the works [3, 4] it was devoted to study the integer ray system (see [3]) by such means that possible to describe moving particles having the integer spin. However, even in the works [2, 5] we actually investigated a systems of ray trajectories which can be characterized as a half-integer ray system [3] by means of which it is possible to describe moving particles having a half-integer spin.

We aim, in the present work, to study a half-integer ray system, two-dimensional and three-dimensional geometric models of motion of the particles having a half-integer spin.

A geometric interpretation of Pauli Principle showing distribution of electrons on energy levels of the atom (such as those described in the physics textbooks [6, 7]) is suggested herein.

The geometric interpretation of the main, azimuthally, magnetic and spin numbers is given in the present work in the form of small angles and distances.

Also, we show a possibility of the existence of the final number of electron shells in the elements of the Mendeleev Periodic System of Chemical Elements. The shells and subshells of the atoms are interpreted as a system of the wave trajectories consisting of direct inclined pieces.

Geometric interpretations of the hydrogen atom and its power levels respectively are separately given in the work as well.

So forth, on the basis of the research of the half-integer ray model, we introduce the step structure of layers in a laminar current of a liquid (such a liquid is described in most textbooks, see [8]).*

*The laminary liquid current was first described long time ago by Newton. The Netwon theory was rechecked many times (see [8]).

Numerical calculations, presented in the present work, as well as those published in [3], were represented by means of three-dimensional tables created in Excel.

For the convenience of readers, reference drawings taken physics textbooks are given in Appendix, while the research part of our publication contains only originally calculated drawings and tables.

1 Half-integer system of eight groups of rays

1.1 Two-dimensional projection of Gaussian (paraxial) rays

In the work [3], we briefly described a paraxial binary (sharing in two) flat system of trajectories. This system consists of groups of rays, in which the rays are inclined under p angles, small to an axis, multiple to the angle γ:

$$p = \left(i + \frac{1}{2}\right)\gamma, \qquad i = 0, \pm1, \pm2, \ldots \tag{1}$$

We called this system of rays: "$(i + 1/2)\gamma$-system" or *half-integer ray system* [3]. We will describe this system in more detail in Fig. 1.

This binary system of rays consists of eight groups of the rays and their links. The rays and links of each of these groups aren't imposed on the rays of other groups, but can cross them.

Branching points of the rays will be spaced from a symmetry axis on small distances of q, multiple to $\frac{1}{2}k$ length:

$$q = \frac{jk}{2}, \qquad j = 0, \pm1, \pm2, \ldots \tag{2}$$

Further, we more precisely will refer to "$(i + 1/2)\gamma$-system" as "$[\, p = (i + 1/2)\gamma,\ q = jk/2\,]$-system".

In this work, as well as in the previous works [2–5] we assume that the rays extend along the branching links; therefore the number of the rays $\mathbb{N}$ can be summarized. We also assume that $\mathbb{K}$ is a number of the links generally $\mathbb{N} \geqslant \mathbb{K}$.

In Fig. 1 (a-d, f-i) eight groups of rays of the aforementioned $[\, p = (i + 1/2)\gamma,\ q = jk/2\,]$-system are shown: $\boldsymbol{K''}$, $\boldsymbol{L''}$, $\boldsymbol{M''}$, $\boldsymbol{N''}$, $\boldsymbol{O''}$, $\boldsymbol{P''}$, $\boldsymbol{Q''}$, $\boldsymbol{R''}$.

This system is placed on a rectangular *coordinate grid*. The size of a cell of a grid has height of $\frac{1}{2}k$ and length of L, $L \gg \frac{1}{2}k$, $L \gg \frac{1}{2}jk$.

Groups in Fig. 1 (a-d) and in Fig. 1 (f-i) are shifted from each other down on the $\frac{1}{2}k$ distance. Groups in Fig. 1 (f-i) are shifted concerning groups in Fig. 1 (a-d) on distance of L.

In Fig. 1 (e) and Fig. 1 (j) the image of groups of the rays $\boldsymbol{K''}$, $\boldsymbol{L''}$, $\boldsymbol{M''}$, $\boldsymbol{N''}$ and $\boldsymbol{O''}$, $\boldsymbol{P''}$, $\boldsymbol{Q''}$, $\boldsymbol{R''}$ respectively, are combined altogether. In Fig. 1 (k) all eight groups of rays are combined together.

1.2 Three-dimensional projection of Gaussian (paraxial) rays

In the work [3] we considered the three-dimensional image of a binary paraxial system of rays in the form of a nonlinear arithmetic parallelepiped: *In a nonlinear arithmetic parallelepiped all numbers are located in the rectangular planes of identical sizes, and these planes are located layer-by-layer one under another since parallelepiped top*.

In this case, the nonlinear arithmetic parallelepiped [3] has a $\mathfrak{N}L$ height, a length of $D = \frac{1}{2}km' + 1$ and a width $\Gamma = \gamma m + 1$, where L, k are distances, while γ is a small angle in the two-dimensional binary ray system (Fig. 1) and at the same time a small distance in a three-dimensional nonlinear parallelepiped [3], and $\mathfrak{N}$, m, m' are natural numbers or zero.

After a large number of passes (iterations) of $\mathfrak{N} \to \infty$ and $\mathfrak{N}L \gg L$, we write down the rule of consecutive filling with numbers of a nonlinear arithmetic parallelepiped as well as in [3]:

$$A = B + C, \tag{3}$$

where

$$A = \begin{pmatrix} \mathfrak{N} \\ p \\ q \end{pmatrix}, \quad B = \begin{pmatrix} \mathfrak{N} - 1 \\ p - 1 \\ q + p - 1 \end{pmatrix}, \quad C = \begin{pmatrix} \mathfrak{N} - 1 \\ p + 1 \\ q + p + 1 \end{pmatrix}.$$

For creation of various types [3] of nonlinear arithmetic parallelepipeds it is necessary to set various additional boundaries and initial conditions.

1.3 Periodic and acyclic trajectories

The system $[\, p = (i + 1/2)\gamma,\ q = jk/2\,]$ of rays generally consists of periodic and acyclic trajectories. In Fig. 2, one of eight groups of the rays of this system are shown for the case of $D = 4k$, $\Gamma = 7\gamma$.

We will set the *first boundary conditions* [3] for number A in formula (3) for nonzero $\mathfrak{N}$-layers:

$$A = 0 \tag{4}$$

for $q = |\, q_{max}|$, where $q_{max} = \frac{1}{2}D$.

Further we will set the *first boundary conditions* for numbers B and C in formula (3) for nonzero $\mathfrak{N}$-layers:

$$B = 0, \qquad C = 0 \tag{5}$$

for $|\, q + p - 1\,| > q_{max}$ and $|\, q + p + 1\,| > q_{max}$ rectively.

We now set the initial conditions [3] for the numbers B and C in formula (3) for the sequence of numbers q of a zero layer ($\mathfrak{N} = 0$):

$$\begin{pmatrix} 0 \\ p \\ q \end{pmatrix} = 1 \tag{6}$$

for $|q| \leqslant q_{max}$ and

$$\begin{pmatrix} 0 \\ p \\ q \end{pmatrix} = 0 \tag{7}$$

for other q.

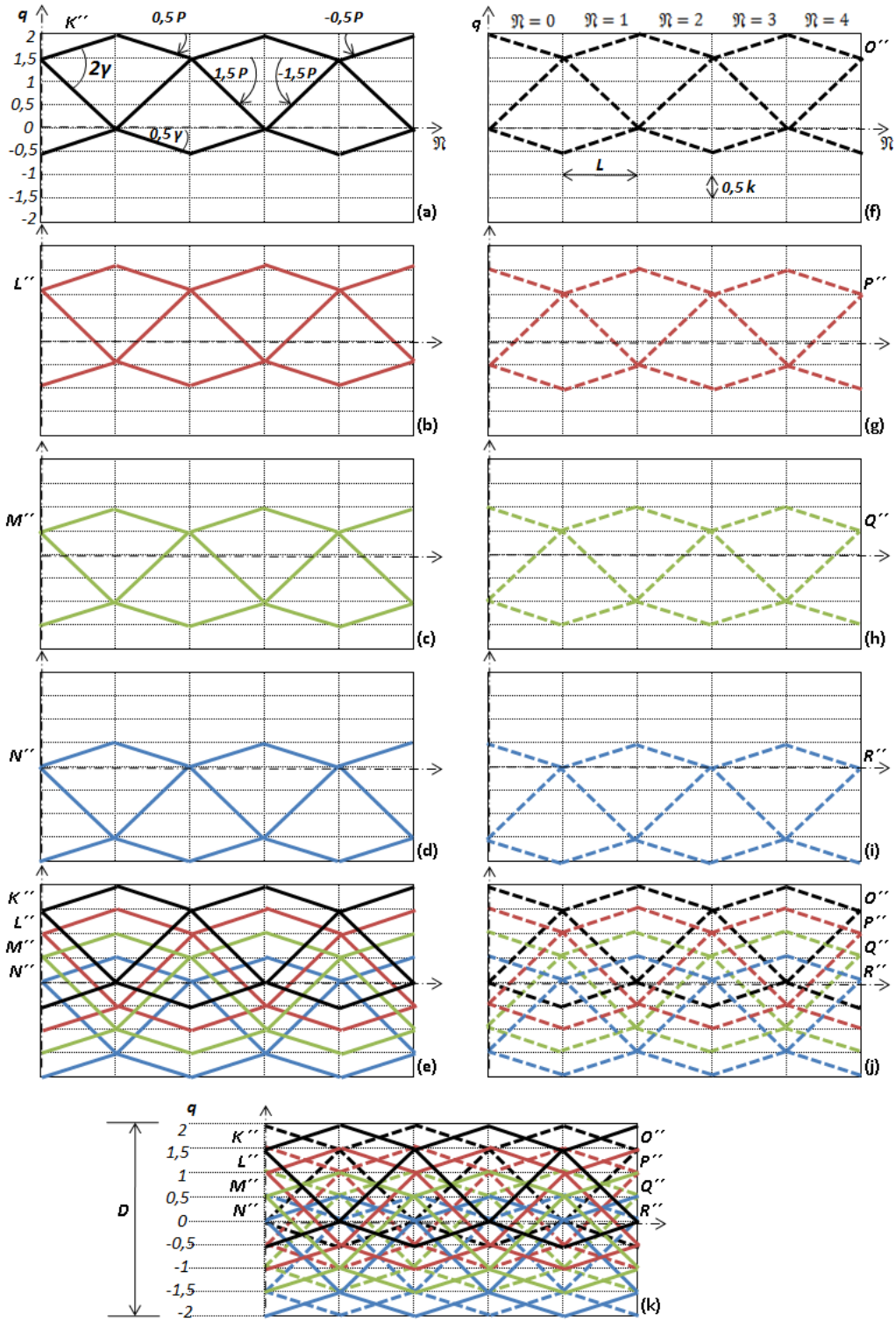

Fig. 1: Periodic trajectories. Eight groups of rays: $\boldsymbol{K''}$, $\boldsymbol{L''}$, $\boldsymbol{M''}$, $\boldsymbol{N''}$, $\boldsymbol{O''}$, $\boldsymbol{P''}$, $\boldsymbol{Q''}$, $\boldsymbol{R''}$ of the $[\, p = (i + 1/2)\,\gamma,\ q = jk/2\,]$ system. $k/2$, and L are the minimum distances on a vertical and a horizontal respectively. $\mathfrak{N}$ is the number of pass of rays (the number of iteration). Dash-dotted lines with arrows showed axes of a coordinate grid in which trajectories are placed.

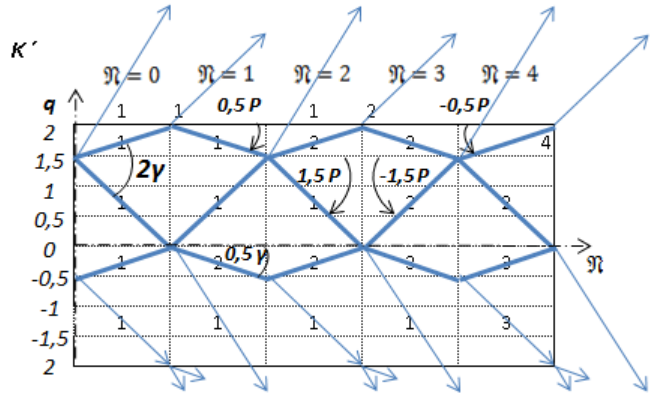

Fig. 2: Periodic and acyclic trajectories. One of eight groups of rays (K' group) of the [$p = (i + 1/2)\,\gamma,\ q = jk/2$] system is shown here. Figures about the shown links illustrate the number of the rays of $\mathbb{N}$ and the summation process of number of the rays extending along the number of the links $\mathbb{K}$ for the first five passes ($\mathfrak{R} = 0 - 4$).

In Fig. 3, calculation formulas (given in MS Excel) of a nonlinear arithmetic parallelepiped (Figs. 1, 2) for $D = 4k$, $\Gamma = 7\gamma$ case for zero, the first and second passes of the ray system, i.e. $\mathfrak{R} = 0, 1, 2$. The calculation was made according to the rule (3) of the consecutive filling with the numbers of an arithmetic rectangle taking into account the boundary (4, 5) and the initial (6, 7) conditions. Three rectangles in Fig. 3 are the layers of a nonlinear arithmetic parallelepiped.

Results of numerical calculation for the first five passes of rays, i.e. $\mathfrak{R} = 0 - 5$ are given in Fig. 4. Five rectangles are layers of a nonlinear arithmetic parallelepiped.

Results of numerical calculations for 32 pass of rays, i.e. for $\mathfrak{R} = 32$ (a) are given in Fig. 5. The envelopes of distribution of number of rays of $\mathcal{K}(q)$ on the section (b) and $\mathcal{K}'(p)$ at the angle (c) are provided.

1.4 Periodic trajectories and step layers in the laminar current of a liquid

In a specific case, the $[p = (i + 1/2)\gamma, q = jk/2]$ system of rays consists only of *periodic trajectories*. Fig. 6 shows one of the eight groups of rays of this system for the case, where $D = 4k$, $\Gamma = 3\gamma$.

In this case, we need to further set special initial and threshold conditions to create the appropriate nonlinear arithmetic parallelepiped [3].

Let's consider here a simple and illustrative (as compared to the description given in [3]) way (an Excel algorithm) of setting special initial and threshold conditions for the parallelepiped that describes the system consisting only of periodic trajectories.

Let's set the *second* threshold conditions for A, B, and C in formula (3) for nonzero $\mathfrak{R}$-layers:

$$A = 0, \tag{8}$$

if $B = 0$ and $C = 0$, and

$$B = 0 \quad \text{and} \quad C = 0, \tag{9}$$

if $A = 0$.

Let's set additional initial conditions for B and C for a zero layer ($\mathfrak{R} = 0$):

$$B = 0 \quad \text{and} \quad C = 0, \tag{10}$$

if $A = 0$.

The offered way (the algorithm) can be easily implemented in numerical calculations in Excel.

At first, we completely fill with units a numerical rectangle of the zero layer ($\mathfrak{R} = 0$) according to formula (6) and formula (7).

Then we fill with numbers a numerical rectangle of the first layer ($\mathfrak{R} = 1$) according to formulas (3 to 5). Some zeroes appear in the first layer.

Then we delete numbers (units) from the cells of the zero-layer rectangle which don't influence cells of the first-layer rectangle.

Then we delete numbers from the cells of the first-layer rectangle which don't depend on the cells of the zero-layer rectangle. We have some new zeroes in the first layer again.

Then again we delete numbers (units) from the cells of the zero-layer rectangle which don't influence the cells of the first-layer rectangle.

And so we repeat this process several times. As a result, we still have cells filled with meaningful numbers which influence other cells, and the cells which depend on other cells. The remained cells describe the $[p = (i + 1/2)\gamma, q = jk/2]$ system consisting only of periodic trajectories.

$[p = (i + 1/2)\gamma, q = jk/2]$ is the system of rays consisting only of periodic trajectories as shown in Fig. 7. The results of calculation of a nonlinear arithmetic parallelepiped (Figs. 1 and 6) are for $D = 4k$, $\Gamma = 3\gamma$ for the zero, first, and second passes of the ray system, i.e. ($\mathfrak{R} = 0, 1, 2$). The calculation was made according to the rule (3) of consecutive filling with numbers of an arithmetic rectangle taking into account the *first* and the *second* threshold (4 and 5; 8 and 9) and initial (6, 7, and 10) conditions, including the algorithm (8 to 10). The three rectangles shown in Fig. 7 are the layers of a nonlinear arithmetic parallelepiped.

Fig. 8 shows the images of layers of a nonlinear arithmetic parallelepiped and a numerical example of calculation of the $[p = (i + 1/2)\gamma,\ q = jk/2]$ system of periodic trajectories (Fig. 7) for $D = 4k$, $\Gamma = 3\gamma$ for zero and the subsequent four passes of rays, i.e. for $\mathfrak{R} = 0 - 4$.

Fig. 9 shows numerical calculations and graphics made in Excel. Numerical calculation for the 32nd pass of rays, i.e. for $\mathfrak{R} = 32$, is given in (a). It also shows the envelopes of distribution of the number of rays of $\mathcal{N}(q)$ on the section (b) and $\mathcal{N}'(p)$ at the angle (c) for $D = 4k$, $\Gamma = 3\gamma$.

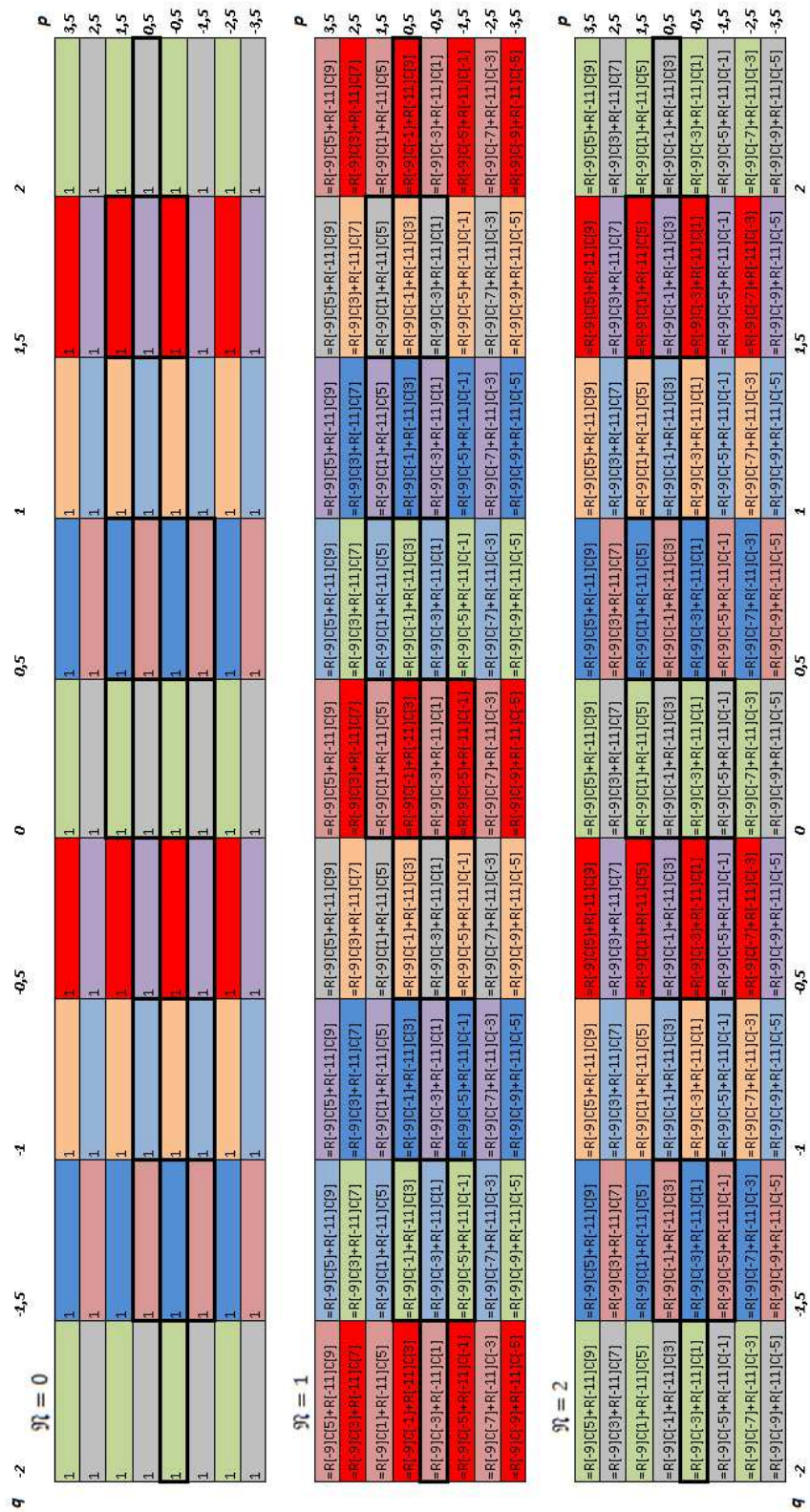

Fig. 3: Calculation of the filling with the numbers of a nonlinear arithmetic parallelepiped for $D = 4k$ case in Excel. The $[\, p = (i + 1/2)\,\gamma,\ q = jk/2\,]$ system of 8 groups of rays of periodic and acyclic trajectories; the first three pass through the rays. Each of the eight groups is marked by an own color.

Fig. 4: Results of numerical calculation in Excel for the first five passes of rays (iterations). Arrows showed dependent cells. Each of eight groups $\boldsymbol{K'}$, $\boldsymbol{L'}$, $\boldsymbol{M'}$, $\boldsymbol{N'}$, $\boldsymbol{O'}$, $\boldsymbol{P'}$, $\boldsymbol{Q'}$, $\boldsymbol{R'}$ of the $[\, p = (i + 1/2)\,\gamma,\ q = jk/2\,]$ system of periodic and acyclic trajectories. Each of the eight groups is marked by an own color.

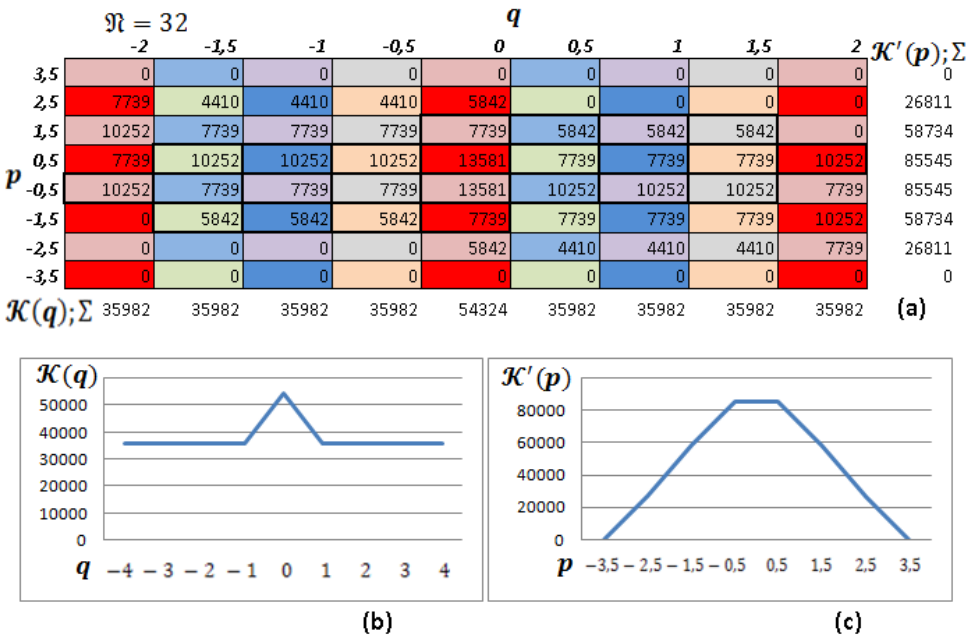

Fig. 5: Results of the numerical calculations for 32 pass of rays, i.e. for $\mathfrak{N} = 32$ for periodic and acyclic trajectories of the considered $[\,p = (i + 1/2)\,\gamma,\ q = jk/2\,]$ system (a); each of the eight groups of the system is marked by an own color; a thick framework in the central part is noted the system of periodic trajectories. The envelopes of distribution of the number of the rays of $\mathcal{K}(q)$ on the section (b) and $\mathcal{K}'(p)$ at the angle (c) are given. We note that for this case, as show our calculations, the form of envelope (b, c) practically doesn't change approximately after the 15th pass.

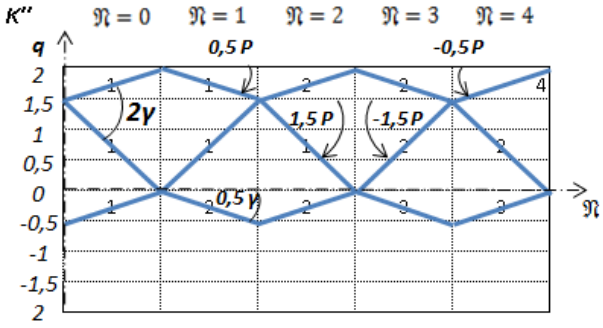

Fig. 6: Periodic trajectories. It shows one of eight groups of rays (***K″*** group) of the $[p = (i + 1/2)\gamma,\ q = jk/2]$ system.

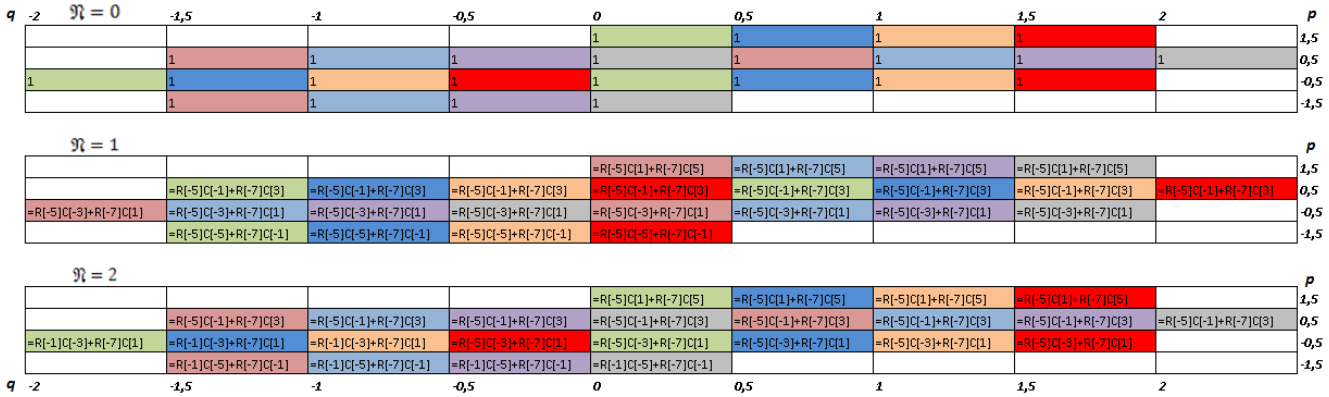

Fig. 7: Calculation of filling with numbers of a nonlinear arithmetic parallelepiped for $D = 4k$ in Excel, and the $[p = (i + 1/2)\gamma,\ q = jk/2]$ system of eight groups of rays of periodic trajectories — the first three passes of rays. Each of the eight groups is marked in a separate color.

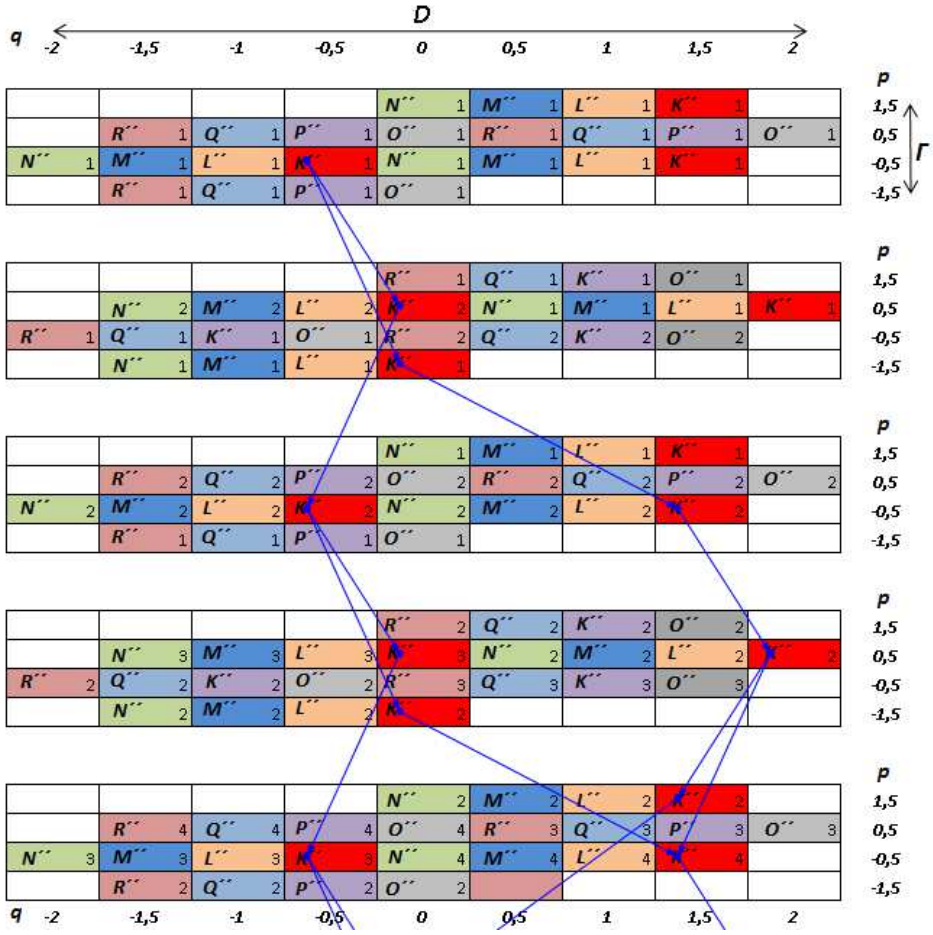

Fig. 8: Results of numerical calculation in Excel for the first five passes of rays (iterations). The arrows point to dependent cells. Each of the eight groups — $\boldsymbol{K''}$, $\boldsymbol{L''}$, $\boldsymbol{M''}$, $\boldsymbol{N''}$, $\boldsymbol{O''}$, $\boldsymbol{P''}$, $\boldsymbol{Q''}$, and $\boldsymbol{R''}$ — of the $[p=(i+1/2)\gamma,\ q=jk/2]$ system of periodic trajectories is marked in a separate color.

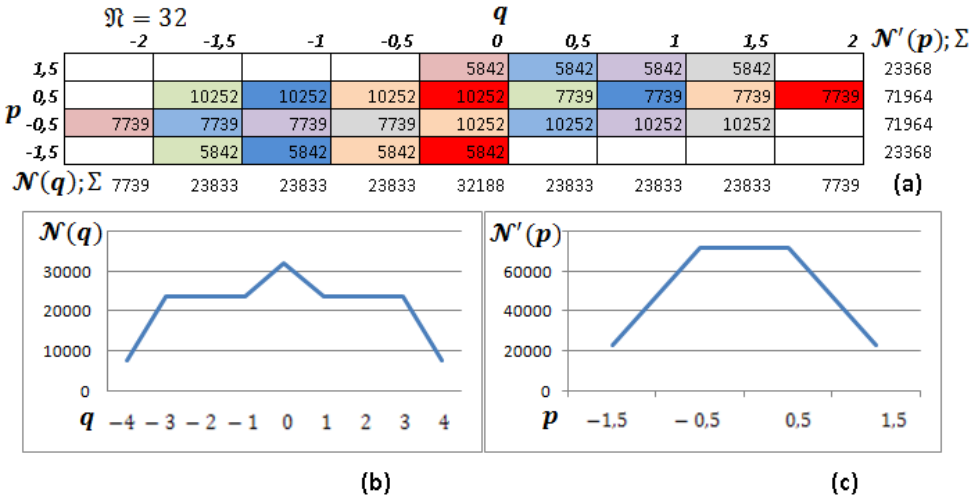

Fig. 9: Results of numerical calculations for the 32nd pass of rays, i.e. for $\Re = 32$, for periodic trajectories of the $[p=(i+1/2)\gamma,\ q=jk/2]$ system (a). Each of the eight groups of the system is marked in separate color. It also shows the envelopes of distribution of number $\mathbb{N}$ of rays of $\mathcal{N}(q)$ on the section (b) and $\mathcal{N}'(p)$ at the angle (c). Note, according to our calculations, in this case, there is virtually no change in the form of the envelope, (b) and (c), approximately after the 15th pass.

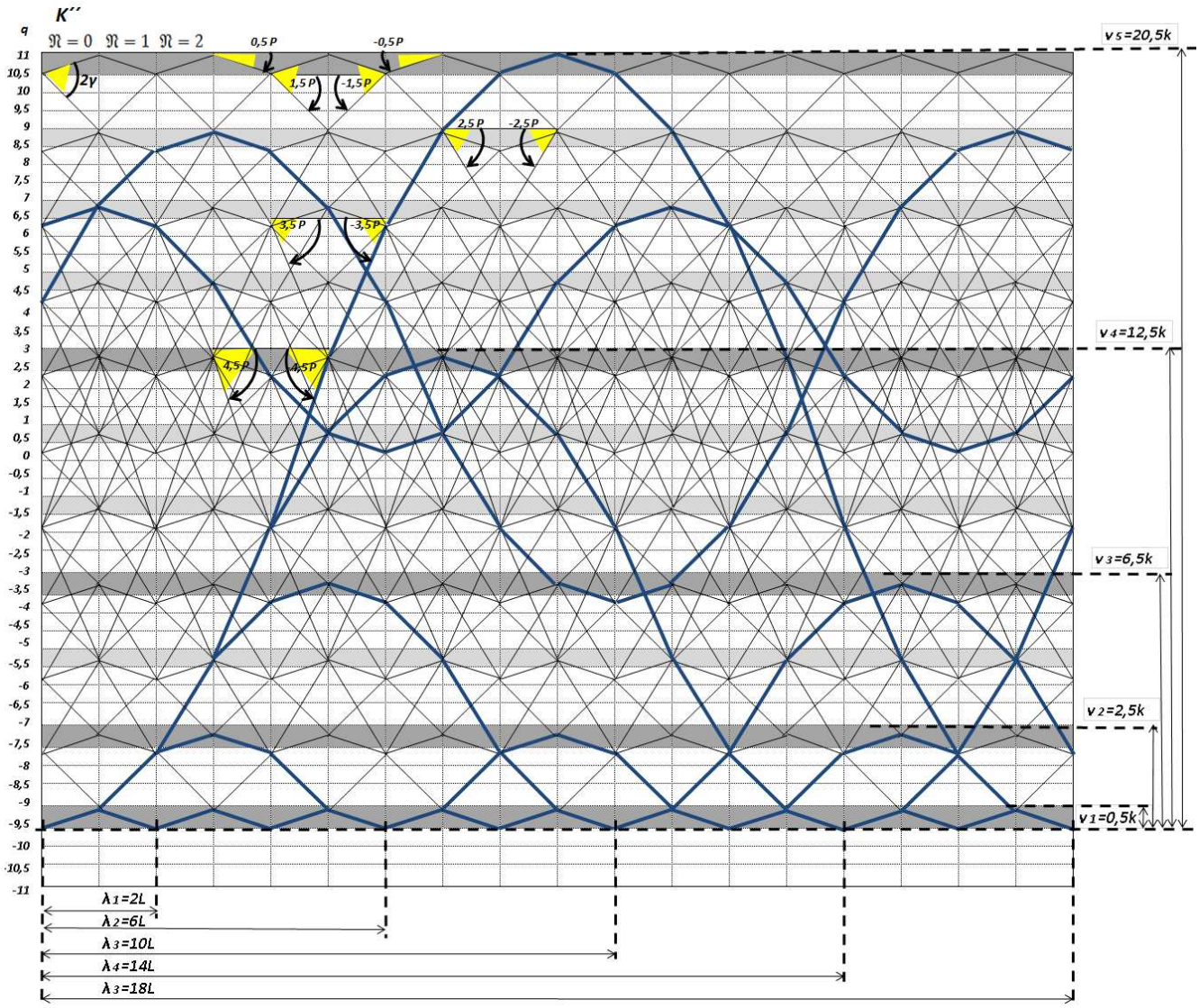

Fig. 10: Periodic (wavy) trajectories. It shows one of the eight groups of rays (***K''*** group) of the $[p=(i+1/2)\gamma,\ q=jk/2]$ system. The group contains 35 links within one pass of $\mathfrak{R}$. Crests and troughs of the "waves" are located (attached) between the horizontals marked in dark color. These horizontals have thickness of $0,5k$ and are located at identical distance of $2k$ from each other.

Fig. 10, similar to Figs. 1 and 6, shows one of the eight groups of rays of the $[p=(i+1/2)\gamma,\ q=jk/2]$ system of rays of periodic trajectories for $D=22k$, $\Gamma=9\gamma$.

In Fig. 10, some of wavy geometric trajectories for the considered $[p=(i+1/2)\gamma, q=jk/2]$ system of rays are shown as heavy lines. Wavy trajectories consist of the links inclined at small angles of $p=(i+1/2)\gamma$. The D size of the binary ray system can accommodate one "wave" or packages of "waves" of different length.

Let's denote the length of wavy trajectories by λ_n. With increasing D this λ_n is growing discretely:

$$\lambda_n = 2\,(2n-1)\,L, \tag{11}$$

where $n=1,2,\ldots$

Let's denote the height of this "wave" by ν_n. The ν_n height is proportional to the squared λ_n length:

$$\nu_n = \frac{\left(n^2-n+\frac{1}{2}\right)}{k} \sim \lambda_n^2 . \tag{12}$$

Wavy trajectories in Fig. 10 can settle down in any part of the *coordinate grid* between horizontals within D.

Fig. 11, similar to Fig. 9, shows numerical calculations and graphs made in Excel. Numerical calculation for the 128th pass of rays, i.e. for $\mathfrak{R}=128$, is given in (a). There are also envelopes of distribution of number $\mathbb{N}$ of rays of $\mathcal{N}(q)$ on the section (b) and $\mathcal{N}'(p)$ at the angle (c) given for the

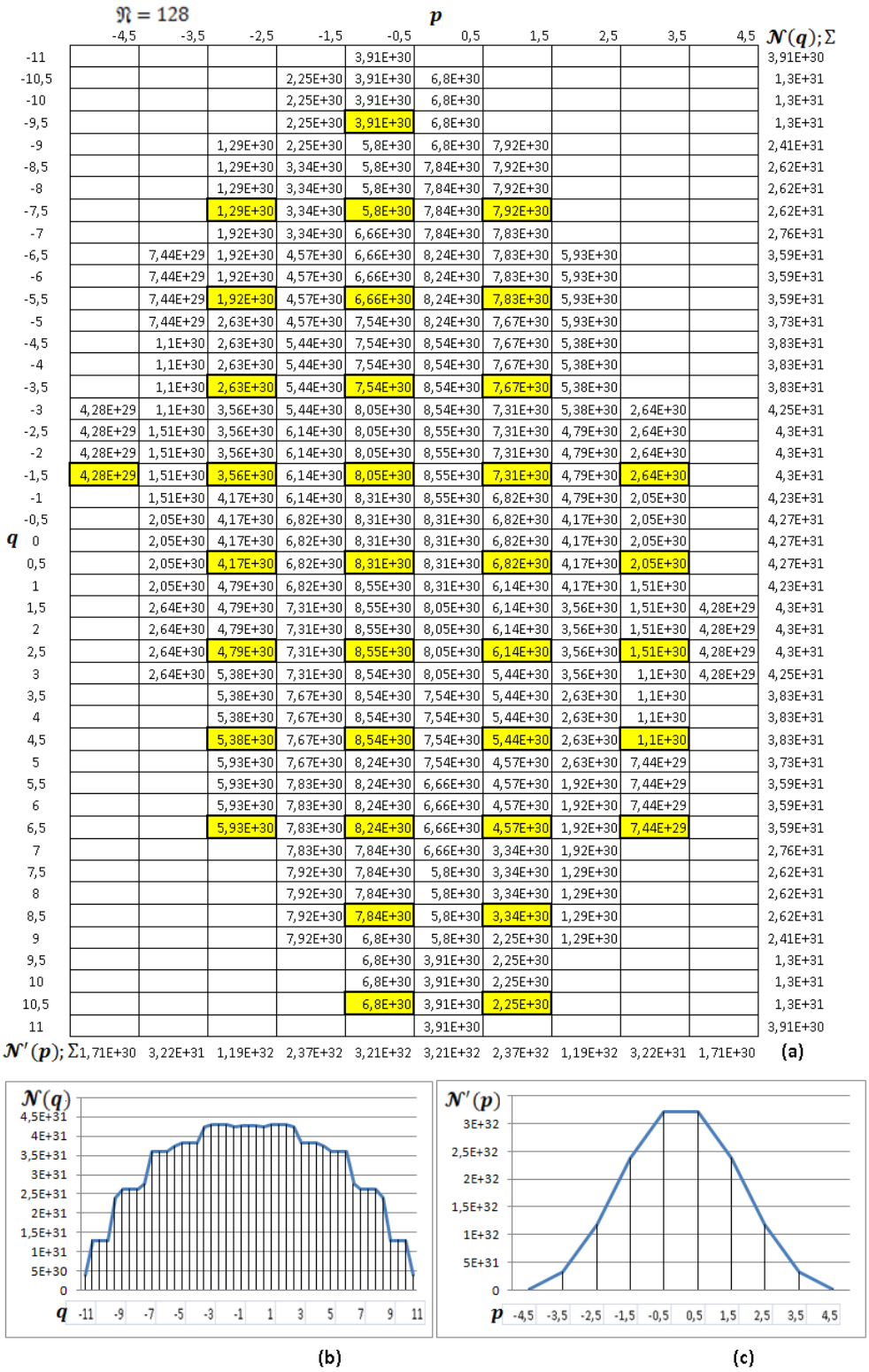

Fig. 11: Results of numerical calculations for the 128th pass of rays, i.e., $\Re = 128$, for eight periodic trajectories of the $[p = (i + 1/2)\gamma,\ q = jk/2]$ system. Thirty five cells (corresponding to 35 rays within one pass in Fig. 10) of one of the eight groups of the system (***K″*** group) are highlighted with the darker color and heavy external borders of the cells (a). The figure also shows the envelopes of distribution of number $\mathbb{N}$ of rays of $\boldsymbol{\mathcal{N}(q)}$ on the section (b) and $\boldsymbol{\mathcal{N}'(p)}$ at the angle (c). Note, according to our calculations, in this case, there is virtually no change in the form of the envelopes, (b) and (c), approximately after the 70th pass.

$[p = (i + 1/2)\gamma,\ q = jk/2]$ system for all the eight groups for $D = 22k$, $\Gamma = 9\gamma$.

The darker cells with heavy external borders shown in Fig. 11 correspond to one of the eight groups of the system (***K″*** group) shown in Fig. 10.

Forms of the envelopes of distribution of the number of rays for the *half-integer system* are similar to the forms of the envelopes for the *integer system* described in [3]. The form of the envelope $\mathcal{N}(q)$ in Fig. 11(b) on the section after a large number of passes (Fig. 11b) is close to a parabola of the fourth degree, and the form of the envelope $\mathcal{N}'(p)$ at the angle (Fig. 11c) is close to Gaussian distribution.

In [3] we noted that the form of the envelope $\mathcal{N}(q)$ on the section for periodic trajectories corresponds to the form of the envelopes of speed distribution at zero pass ($\mathfrak{N} = 0$) and volume distribution after a large number of passes ($\mathfrak{N} \to \infty$) of liquid in pipe section at laminar flow.

In Fig. 11(b) we can see that the envelope $\mathcal{N}(q)$ has a stepped structure compared to the more smooth form of the envelope $\mathcal{N}'(p)$ (Fig. 11c). Similar results were received from numerical calculations for the integer system in [3], but the half-integer model gives the more accurate image of the "steps" compared to the integer model.

It can be assumed that such a stepped structure of the envelope $\mathcal{N}(q)$ explains the existence of layers of final thickness in liquid at laminar flow [8]. The speed and volume of liquid do not change within each of these layers of a certain final thickness.

2 Gaussian (paraxial) rays and Pauli Principle

2.1 Angles, distances and quantum system

Pauli Principle [6, 7] is correct for electrons and other particles with half-integer spin in a quantum system.

The condition of each electron in an atom is characterized by four quantum numbers [6, 7]:

$$\left.\begin{array}{rll} \text{Principal} & n & (n = 1, 2, 3, \dots) \\ \text{Azimuthal} & l & (l = 0, 1, 2, \dots, n-1) \\ \text{Magnetic} & m_l & (m_l = -l, \dots, -1, 0, +1, \dots, +l) \\ \text{Spin} & m_s & \left(m_s = +\tfrac{1}{2},\ -\tfrac{1}{2}\right) \end{array}\right\}. \quad (13)$$

Fig. 23 of the Appendix illustrates an example from [6] of spatial quantization.

In monographs [6] and [7] the spin is also denoted by one letter "s":

$$m_s = s = \pm\frac{1}{2}. \quad (13a)$$

According to Pauli Principle in a quantum system, for example in an atom, there can't be two electrons possessing identical quantum numbers: n, l, m_l, m_s. That is, two electrons cannot be in the same state simultaneously. No more than $2n^2$ electrons can be in a state with n value in an atom [6, 7]. If

$$\left.\begin{array}{ll} n = 1 & \text{there can be 2 electrons} \\ n = 2 & \text{there can be 8 electrons} \\ n = 3 & \text{there can be 18 electrons} \\ \dots\dots\dots\dots\dots\dots\dots\dots & \end{array}\right\}. \quad (14)$$

Electrons having identical value of the quantum number n form a *shell*. Shells consist of *subshells*, differing in value of the quantum number l.

Shells are denoted by characters according to value of n [6] and [7]:

$$\left.\begin{array}{rl} \text{Value of } n & 1\ 2\ 3\ 4\ 5\ 6\ 7\dots \\ \text{Designation of the shell} & K\ L\ M\ N\ O\ P\ Q\dots \end{array}\right\}. \quad (15)$$

The electron which is in condition of $l = 0$ is called an s electron, $l = 1$ — p electron, $l = 2$ — d electron, $l = 3$ — f electron, followed by g, h, etc. alphabetically. The value of the principal quantum number n is specified before the symbol of the azimuthal quantum number l [7].

The division of possible conditions of an electron in an atom into shells and subshells [7] is presented in the form of a *periodic table of conditions* of an electron (see Fig. 24 of the Appendix).

The process of building electron shells [7] (according to Pauli Principle) of the first 36 elements of the Mendeleev Periodic System is presented in the form of a *periodic table of elements* (see Fig. 25 of the Appendix).

Now let's give an algorithm of creation of another specific case of the binary $[p = (i + 1/2)\gamma,\ q = jk/2]$ system of rays of periodic trajectories (considered in Section 1.4). We will begin with the minimum quantity of rays consistently passing to the more complicated configurations of the system. Thus, we will compare the properties of our system to the data provided in periodic tables in Figs. 24 and 25 of the Appendix.

We accept that, for our paraxial beams, all the angles of γn, are small and multiple to the small angle of γ, and the small distance of k is as follows:

$$k \approx \gamma L. \quad (16)$$

For perfect correspondence between our geometric constructions and expressions (13 and 13a), including the data provided in periodic tables in Figs. 24 and 25 of the Appendix, we will enter the following assumptions:

$$\text{Principal number} \quad n \sim \gamma i \sim kj, \quad (n = 1, 2, \dots;\ i = 1, 2, \dots;\ j = 1, 2, \dots), \quad (17)$$

$$\text{Azimuthal number} \quad l = n - 1 \sim \gamma(n-1), \quad (18)$$

$$\text{Magnetic number} \quad m_l = \pm l \sim \pm k(n-1), \quad (19)$$

$$\text{Spin number} \quad s = \pm\frac{1}{2} \sim \pm\frac{\gamma}{2} \text{ and } m_s = \pm\frac{1}{2} \sim \pm\frac{k}{2}. \quad (20)$$

Electron shell	n	l	m_l	m_s	Sub shell
K	1	0	0	↑↓	$K(1s)$

(a)

Element	K	L		M		
	$1s$	$2s$	$2p$	$3s$	$3p$	$3d$
1 H	1	-	-	-	-	-
2 He	2	-	-	-	-	-

(b)

Fig. 12: K shell and the first parts of the periodic tables of (a) conditions of an electron, and (b) elements.

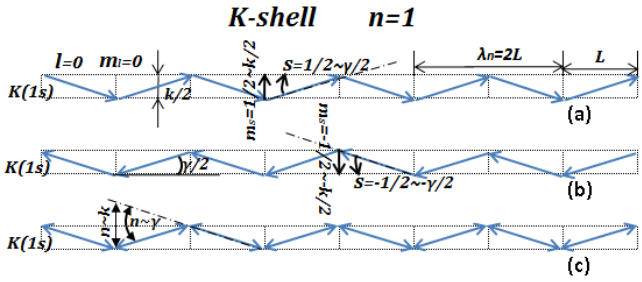

Fig. 13: One of the eight groups of rays of the $\{p=(i+1/2)\gamma\ ,\ q=jk/2\}$ subsystem of periodic trajectories, and K shell. (a) and (b) correspond to an atom of hydrogen, (c) — to an atom of helium. $n=1\sim\gamma\sim k$, $l=0$, $m_l=0$, $s=\pm1/2\sim\pm\gamma/2$, $m_s=\pm1/2\sim\pm k/2$. Dash-dotted lines show axes from which sizes of angles and distances are counted.

Electron shell	n	l	m_l	m_s	Sub shell
L	2	0	0	↑↓	$L_1(2s)$
		1	−1	↑↓	$L_2(2p)$
			0	↑↓	
			+1	↑↓	

(a)

Element	K	L		M		
	$1s$	$2s$	$2p$	$3s$	$3p$	$3d$
3 Li	2	1	-	-	-	-
4 Be	2	2	-	-	-	-
5 B	2	2	1	-	-	-
6 C	2	2	2	-	-	-
7 N	2	2	3	-	-	-
8 O	2	2	4	-	-	-
9 F	2	2	5	-	-	-
10 Ne	2	2	6	-	-	-

(b)

Fig. 14: L shell and the second parts of the periodic tables of (a) conditions of an electron, and (b) elements.

2.2 Periodic tables and geometrical constructions

2.2.1 Creation of the first shell of a quantum system

Fig. 12(a) shows the first (top) part of the periodic table of conditions of an electron (Fig. 24 of the Appendix) describing the first shell of K. Fig. 12(b) shows the top part of the periodic table of elements (Fig. 25 of the Appendix) describing the first two elements:

Fig. 13 shows one of the eight similar to (Figs. 1 and 6) groups of rays of the $[p=(i+1/2)\gamma,\ q=jk/2]$ *system* of periodic trajectories.

These trajectories correspond to the first electron shell of K shown in Fig. 12(a).

To be specific, let's call this *system* of periodic trajectories an $\{p=(i+1/2)\gamma,\ q=jk/2\}$ *subsystem of periodic trajectories* of the $[p=(i+1/2)\gamma,\ q=jk/2]$ *system* of periodic trajectories.

Fig. 13 (a and b) shows two trajectories with opposite (↑ symbol and ↓ symbol) orientation of a spin (one 1s electron). These trajectories correspond to an atom of hydrogen with random orientation of the spin (Fig. 12b).

Fig. 13c shows a trajectory with anti-parallel (↑↓ symbol) spin orientation (two 1s electrons). This trajectory corresponds to an atom of helium (Fig. 12b).

The atom of helium is closing filling of the K shell.

2.2.2 Creation of the second shell of a quantum system

Fig. 14 (a) shows the second part of the periodic table of conditions of an electron (Fig. 24 of the Appendix) describing the second cover of L. Fig. 14 (b) shows the second part of the periodic table of elements (Fig. 25 of the Appendix) describing the elements number three to ten.

Fig. 15 shows one of the eight similar to (Figs. 1 and 6) groups of rays of the $\{p=(i+1/2)\gamma,\ q=jk/2\}$ subsystem of periodic trajectories. These trajectories correspond to the second electron shell of L in Fig. 14 (a and b).

Fig. 15 (a) shows the L shell in the form of a periodic trajectory consisting of the subshell L_1 (one 1s electron) for Li. The form of this shell is the same as in Fig. 13(a) or in Fig. 13(b). The form of K shell for Li is the same as in Fig. 13(c).

Fig. 15 (b, c, d, e, f, g, and h) shows the L shells ($L_1(2s)$ and $L_2(2p)$ subshells) for Be, B, C, N, O, F, and Ne respectively (Fig. 14a and b):

The K shell for these elements is the same as that for Li (see Fig. 13c).

The K shell and L shell of the elements (Figs. 12 to 15) can settle down in our geometric model similar to arrangement of $\boldsymbol{K''}$ group of rays and $\boldsymbol{L''}$ group of rays respectively (Figs. 1, 7 and 8).

The atom of Ne is closing filling the L shell.

2.2.3 Creation of the third and fourth shells of a quantum system

Geometric schemes of Pauli Principle and elements of periodic table are further constructed in compliance with the above algorithm. Therefore, we will confine ourselves to giving specific examples.

The second part of the periodic table of conditions of an electron (Fig. 24 of the Appendix) describing the third M

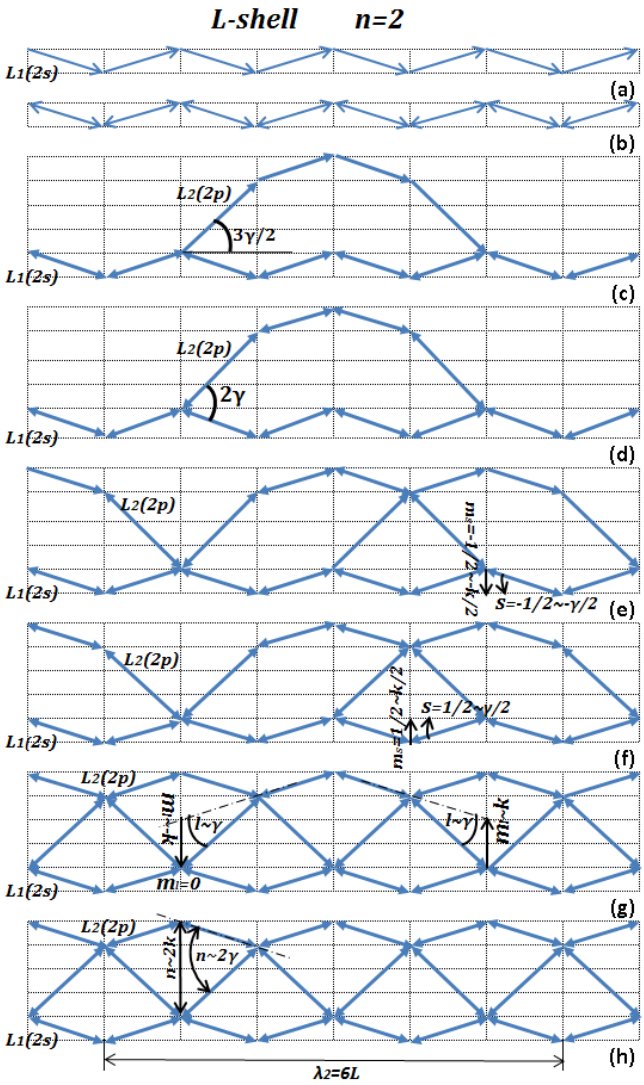

Fig. 15: One of the eight groups of rays of the $\{p=(i+1/2)\gamma,\ q=jk/2\}$ subsystem of periodic trajectories and consecutive process of filling L shell according to Fig. 14 (a and b). Designations of quantum numbers are similar to these in Fig. 13. Dash-dotted lines show axes from which sizes of angles and distances are counted.

shell and the fourth N shell is given in Fig. 16(a). Eight elements of the third part of the periodic table of elements (see Fig. 25 of the Appendix) are given in Fig. 14(b) on a selective basis (see Fig. 16).

Shells of K and L (Fig. 16a) for all the elements shown in Fig. 16(b) are the same as in Fig. 13(c) and Fig. 15(h).

The shell of M ($3s$ subshell) for Na is the same as in Fig. 13(a) or Fig. 13(b).

The subshells $3s$ and $3p$ of the shell of M for Ar and K (Fig. 16b) has the same forms as shown in Fig. 15(h).

The shell of N ($4s$ subshell) for K and Cr (Fig. 16b) is the same as in Fig. 13(a), Fig. 13(b), and Fig. 15(a).

The shell of N ($4s$ subshell) for Sc and Ni is the same as

Electron shell	n	l	m_l	m_s	Sub shell
M	3	0	0	↑↓	$M_1(3s)$
		1	−1	↑↓	$M_2(3p)$
			0	↑↓	
			+1	↑↓	
		2	−2	↑↓	$M_3(3d)$
			−1	↑↓	
			0	↑↓	
			+1	↑↓	
			+2	↑↓	
N	4	0	0	↑↓	$N_1(4s)$
		1	−1	↑↓	$N_2(4p)$
			0	↑↓	
			+1	↑↓	
		2	−2	↑↓	$N_3(4d)$
			−1	↑↓	
			0	↑↓	
			+1	↑↓	
			+2	↑↓	
		3	−3	↑↓	$N_4(4f)$
			−2	↑↓	
			−1	↑↓	
			0	↑↓	
			+1	↑↓	
			+2	↑↓	
			+3	↑↓	

(a)

Element	K	L		M			N	
	$1s$	$2s$	$2p$	$3s$	$3p$	$3d$	$4s$	$4p$
11 Na	2	2	6	1	-	-	-	-
18 Ar	2	2	6	2	6	-	-	-
19 K	2	2	6	2	6	-	1	-
21 Sc	2	2	6	2	6	1	2	-
24 Cr	2	2	6	2	6	5	1	-
28 Ni	2	2	6	2	6	8	2	-
31 Ga	2	2	6	2	6	10	2	1
36 Kr	2	2	6	2	6	10	2	6

(b)

Fig. 16: Shells of M and N, and the third parts of the periodic tables of (a) conditions of an electron, and (b) eight elements.

in Fig. 13(c) and Fig. 15(b).

The shell of N (subshells $4s$ and $4p$) for Ga is the same as in Fig. 15(c).

The shell of N (subshells $4s$ and $4p$) for Kr is the same as in Fig. 15(h).

Fig. 17 (a, b, c, and d) shows the shell of M (subshells 3s, 3p, and 3d) for (Sc), (Cr), (Ni), (Ga and Kr) respectively (Fig. 16a and b):

Three or four shells of K, L, M, and N (Figs. 13 to 17) for the elements can settle down in our geometric model similar to the arrangement of groups of rays of $\boldsymbol{K''}$, $\boldsymbol{L''}$, $\boldsymbol{M''}$, and $\boldsymbol{N''}$ in Figs. 1, 7 and 8.

The geometric schemes of Pauli Principle and elements of the periodic table are also further created in compliance with the above algorithm.

However, our geometric model similar to (Figs. 1 and 6) of groups of rays of the $\{p=(i+1/2)\gamma,\ q=jk/2\}$ subsystem of periodic trajectories consists only of eight groups of rays. Therefore, while remaining within the offered model, it is possible to assume that the number of shells of an atom is no more than eight either. If we continue increasing the number of ray groups to more than eight, the rays will overlap, and the shells will merge.

Thus, if the number of shells does not exceed eight, the total number of elements of the periodic system (14) cannot exceed 128.

Deviations from the sequence of filling the periodic system (e.g., for the elements such as K, etc.) (Fig. 16b) hypothetically reduce (or increase) the total number of elements of the periodic system.

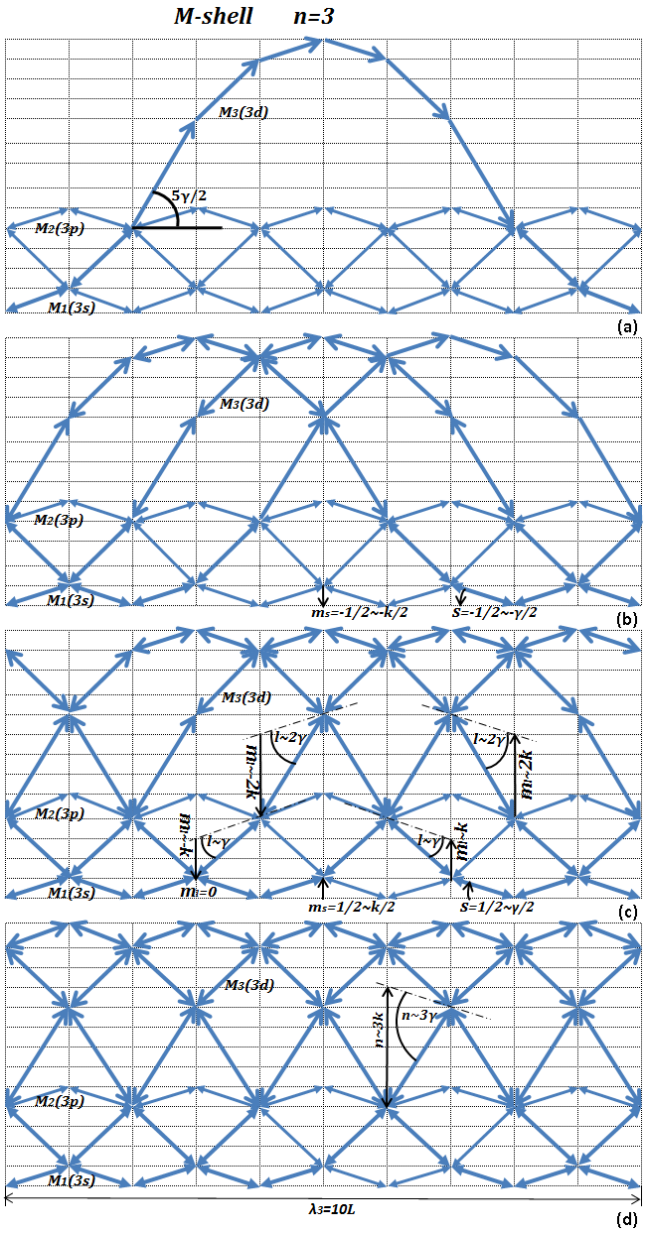

Fig. 17: One of the eight groups of rays of $\{p = (i + 1/2)\gamma,\ q = jk/2\}$ subsystem of periodic trajectories and consecutive process of filling M and N shells for eight elements according to Fig. 16 (a and b). Designations of quantum numbers are similar to these in Figs. 13 and 15. Dash-dotted lines show axes from which sizes of angles and distances are counted.

2.2.4 Pauli Principle and the geometric system of the hydrogen atom

Monographs on quantum mechanics [6] and [7] consider the simplest quantum mechanical system of an atom of hydrogen (Figs. 26 to 28 of the Appendix). Let's make a review of this example too.

Fig. 18 shows the fifth shell of O:

Electron shell	n	ℓ	m_ℓ	m_s	Sub shell
O	5	0	0	⇅	$O_1(5s)$
		1	−1	⇅	$O_2(5p)$
			0	⇅	
			+1	⇅	
		2	−2	⇅	$O_3(5d)$
			−1	⇅	
			0	⇅	
			+1	⇅	
			+2	⇅	
		3	−3	⇅	$O_4(5f)$
			−2	⇅	
			−1	⇅	
			0	⇅	
			+1	⇅	
			+2	⇅	
			+3	⇅	
		5	−4	⇅	$O_5(6g)$
			−3	⇅	
			−2	⇅	
			−1	⇅	
			0	⇅	
			−1	⇅	
			−2	⇅	
			−3	⇅	
			−4	⇅	

Fig. 18: Shell of O of the periodic table of conditions of an electron.

Fig. 19 shows one of the eight groups similar to (Figs. 1 and 6) of rays of the $\{p = (i + 1/2)\gamma,\ q = jk/2\}$ subsystem of periodic trajectories. These trajectories correspond to the fifth electron shell of O shown in Fig. 18. In this example, the O shell is filled completely and contains five subshells.

The $\{p = (i + 1/2)\gamma,\ q = jk/2\}$ subsystem for an atom of hydrogen can be constructed geometrically in accordance with Pauli Principle and similar to the construction method described in previous Sections.

The $\{p = (i + 1/2)\gamma,\ q = jk/2\}$ subsystem shown in Fig. 19 is in many respects similar to the $[p = (i + 1/2)\gamma,\ q = jk/2]$ system in Fig. 10, but contains the smaller quantity of rays and the smaller quantity of the wavy trajectories consisting of these rays.

The wavy trajectories shown in Fig. 19 settle down in the lower part of the coordinate grid and are "attached" to the lower horizontal unlike the wavy trajectories in Fig. 10, which can settle down in any part of the coordinate grid within D size.

In principle, the creation of the O shell in Fig. 19 does not differ from creation of other shells shown in Figs. 13, 15, 17.

Upon comparison of angles multiple p in Fig. 10 and multiple n in Fig. 19, it can be seen that the relationship between these angles is as follows:

$$n \sim |p| + \frac{1}{2}\,. \tag{21}$$

In Fig. 19 we illustrated the allowed quantum transitions [6] and [7]:

$$\Delta n = \pm 1 \quad \text{and} \quad \Delta l = \pm 1 \tag{22}$$

in the form of angles, but not distances as in Figs. 26 to 28 of the Appendix. However, considering ratios (16 to 20) for small angles, sizes (22) can be illustrated (in principle) in the form of distances as well, since:

$$\Delta n = \pm 1 \sim \pm\gamma \sim \pm k \quad \text{and} \quad \Delta l = \pm 1 \sim \pm\gamma \sim \pm k\,. \tag{23}$$

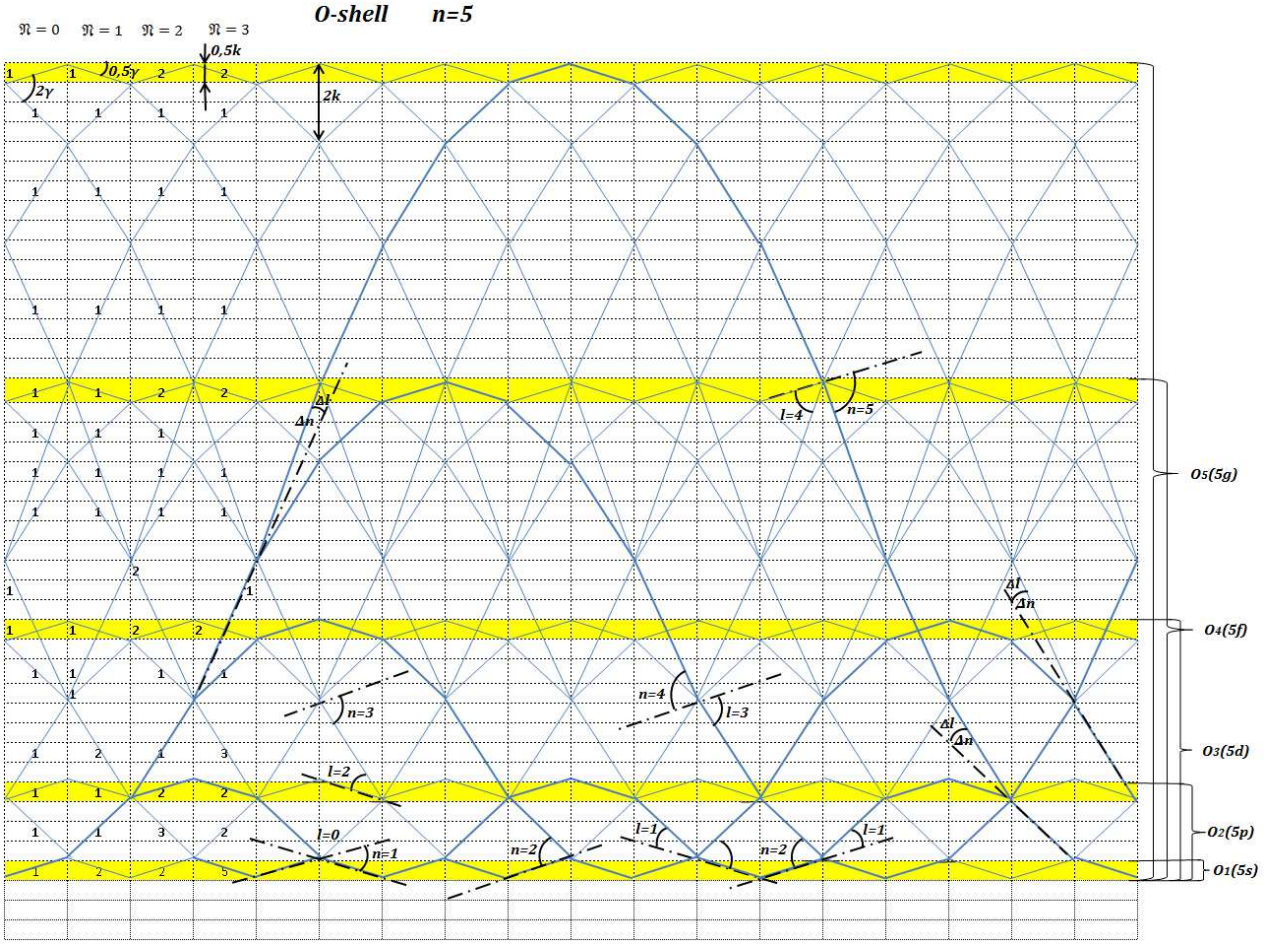

Fig. 19: One of the eight groups of rays of the considered $\{p = (i + 1/2)\gamma,\ q = jk/2\}$ subsystem of periodic (wavy) trajectories. The shell of O is completely filled according to Fig. 18. This group contains 15 links ($\mathbb{K} = 15$) within one pass of $\mathfrak{R}$. Troughs of the "waves" are located ("attached to") only on the lower horizontal, while crests of the "waves" are located ("attached to") on several higher horizontals. The horizontals are marked in separate color. These horizontals have thickness of $0,5k$ and are located at the increasing distance of $2kn$ from each other from the bottom upwards. The designations of the quantum numbers are similar to these in Figs. 13 to 17. Dash-dotted lines show axes from which sizes of angles and distances are counted. Quantum transitions $\Delta n = \pm 1$ and $\Delta l = \pm 1$ are shown in the form of angles between dash-dotted lines. On the right, location of five subshells is indicated by curly braces. Figures put next to links (for $\mathfrak{R} = 0 - 3$) show the number of rays of $\mathbb{N}$ and the process of summation of rays spreading along the links.

Fig. 20 gives images of layers of a nonlinear arithmetic parallelepiped and a numerical example of calculation of O shell of the $\{p = (i + 1/2)\gamma,\ q = jk/2\}$ subsystem of periodic trajectories (Fig. 19) for zero and the subsequent three passes of rays, i.e. for $\mathfrak{R} = 0 - 4$. Four rectangles shown in Fig. 20 are the layers of the nonlinear arithmetic parallelepiped.

n and l values are given on the right in Fig. 20 with taking into account the ratios (18 and 21).

The calculation was made according to the rule (3) of consecutive filling of an arithmetic rectangle with numbers taking into account the *first* and the *second* threshold (4 and 5; 8 and 9) and initial (6, 7, and 10) conditions, including the algorithm (8 to 10).

Each of the four layers of the arithmetic parallelepiped shown in Fig. 20 is similar to the layer represented in Fig. 11, but there are differences as well. Therefore, it is necessary to set the *third* threshold conditions. We took these conditions for our example directly from Fig. 19. As this approach is illustrative for us, the total number of rays is not that big. All the wavy trajectories are "attached" to the lower horizontal. The *third* threshold conditions can be set in other illustrative ways, e.g., by means of special nomograms.

Thus, we made calculations in Excel according to expression (3). The appropriate formulas for the respective p and q can be taken from Fig. 3 or Fig. 7, for example.

Fig. 21 shows results of numerical calculations of layers of the nonlinear arithmetic parallelepiped for O shell (Fig. 20) in Excel are given in the form of envelopes of distribution of the ray number. In a, c, e, g, and i (the left column), you can see the envelopes of distribution of the ray number at the an-

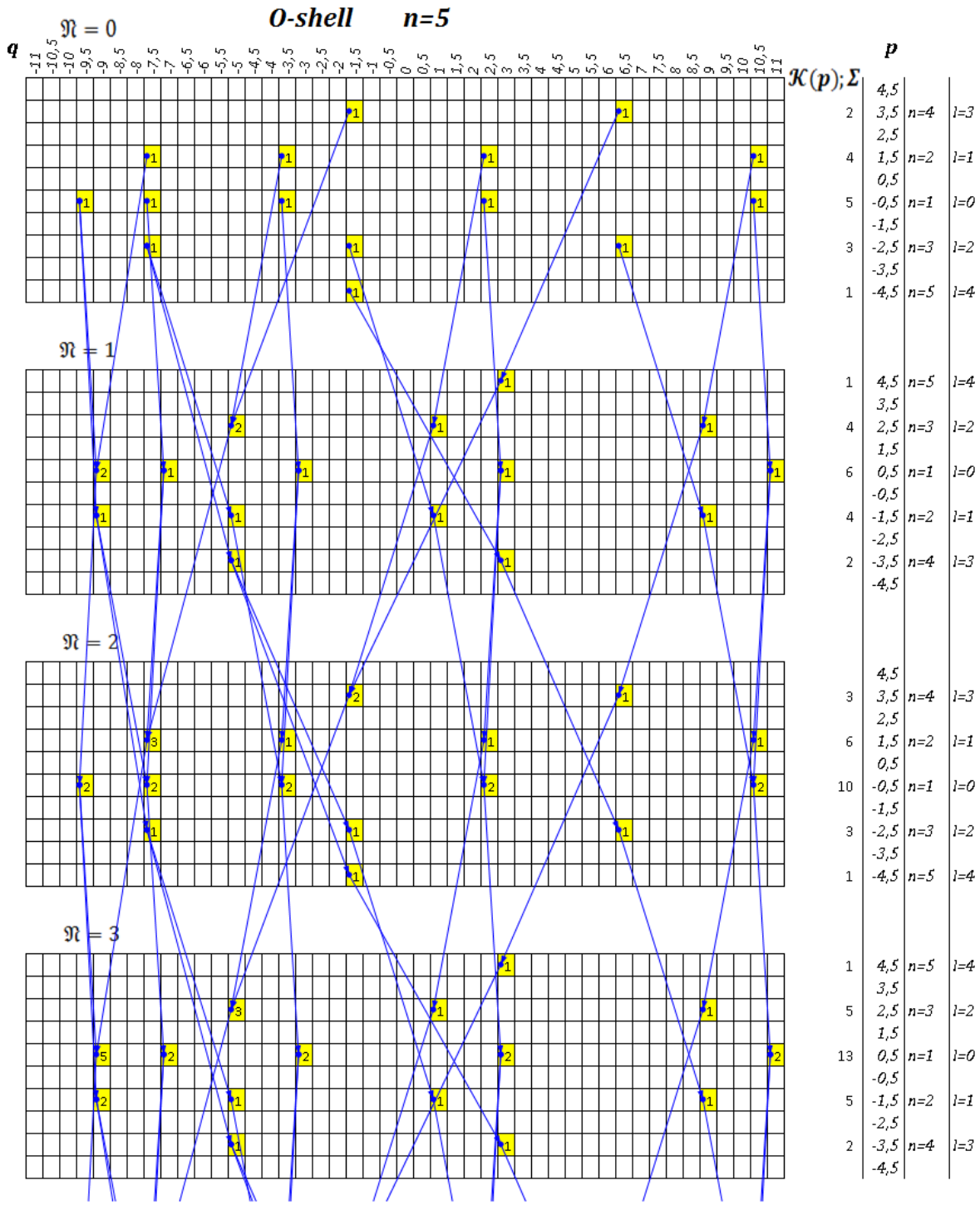

Fig. 20: One of the eight groups of rays of the $\{p=(i+1/2)\gamma,\ q=jk/2\}$ subsystem of periodic (wavy) trajectories. The O shell is completely filled according to Figs. 18 and 19. Numerical calculation in Excel was made for the first four passes of rays (iterations). Arrows show dependent cells. Fifteen highlighted cells within one pass of $\mathfrak{N}$ (one layer of a parallelepiped) correspond to fifteen links ($\mathbb{K}=15$) within one pass of $\mathfrak{N}$ shown in Fig. 19. Figures in the highlighted cells correspond to the number of rays of $\mathbb{N}$ extending along the links of $\mathbb{K}$.

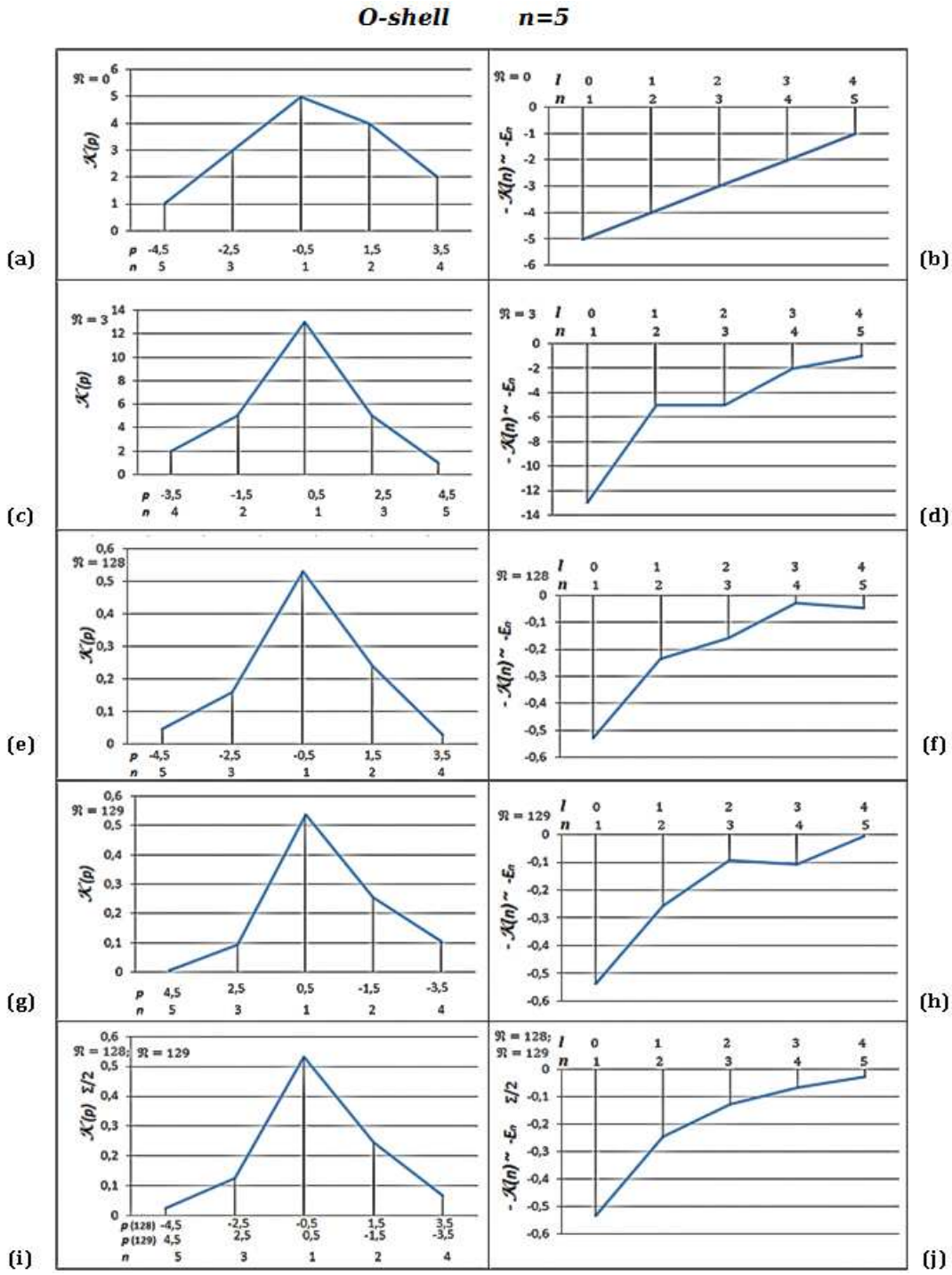

Fig. 21: This Figure shows the results of numerical calculations (Fig. 20) for $\mathfrak{N} = 0$, 32, 128, 129 (a to h) of the completely filled O shell for the $\{p = (i + 1/2)\gamma,\ q = jk/2\}$ subsystem of periodic (wavy) trajectories. The envelopes of distribution of the ray number at the angle $\mathcal{K}(p)$ are given in the left column. The envelopes of distribution of the ray number at the angle $\mathcal{K}(n)$ are given in the right column. The shared envelopes for $\mathfrak{N} = 128, 129$ are given in (i and j). Graphs (e to j) are shown in a normalized form. Note, according to our calculations, in this case, there is virtually no change in the form of the envelope approximately after the 70th pass.

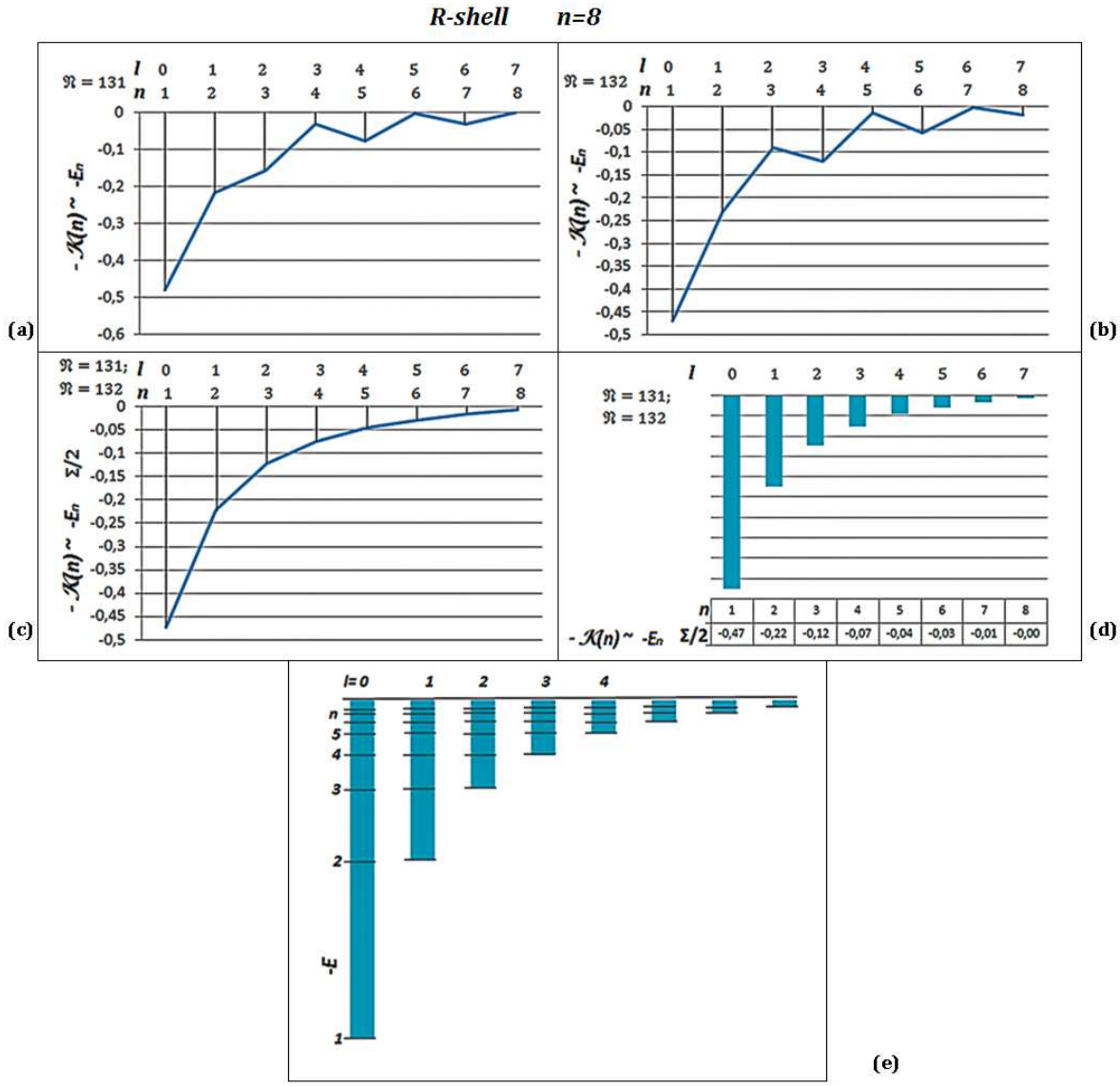

Fig. 22: This Figure shows results of numerical calculations of the completely filled 8th shell of R for the $\{p=(i+1/2)\gamma,\ q=jk/2\}$ subsystem of periodic (wavy) trajectories. In (a) and (b), you can see the envelopes of distribution of the ray number at the angle $\mathcal{K}(n)$ for passes $\mathfrak{R}=131,132$. Results of joint calculations for passes $\mathfrak{R}=131$ and $\mathfrak{R}=132$ in the form of the envelope are given in (c) and these in a form of histograms are given in (d) and (e). Envelopes (a to c) and histograms (d and e) are presented in the normalized form. Note, according to our calculations, in this case, there is virtually no change in the form of envelopes approximately after the 100th pass ($\mathfrak{R}=100$). The histogram in (e) is similar to Fig. 28 of the Appendix.

gle $\mathcal{K}(p)$. In b, d, f, h, and j (the right column), you can see the envelopes of distribution of the ray number at the angle $\mathcal{K}(n)$ (taking into account expression (21). The results of numerical calculation for zero pass of rays, i.e. for $\mathfrak{R}=0$, are given in a and b; for $\mathfrak{R}=3$ — in c and d; for $\mathfrak{R}=128$ — in e and f; and for $\mathfrak{R}=129$ — in g and h.

Shared graphs of $\mathcal{K}(p)$ and $\mathcal{K}(n)$ for $\mathfrak{R}=128$ and $\mathfrak{R}=129$ are given in i and j, namely:

$$\mathcal{K}(p)_{\mathfrak{R}=128,\ \mathfrak{R}=129}=\frac{1}{2}\left[\mathcal{K}(p)_{\mathfrak{R}=128}+\mathcal{K}(p)_{\mathfrak{R}=129}\right] \quad (24)$$

and

$$\mathcal{K}(n)_{\mathfrak{R}=128,\ \mathfrak{R}=129}=\frac{1}{2}\left[\mathcal{K}(n)_{\mathfrak{R}=128}+\mathcal{K}(n)_{\mathfrak{R}=129}\right], \quad (25)$$

where $\mathcal{K}(p)$ and $\mathcal{K}(n)$ are in the normalized form.

In this work, like in previous works [2] to [5], we assume that the number of rays of $\mathbb{N}$ extending along the number of multiplicative links of $\mathbb{K}$ is proportionate to energy. For negative energy of an electron — E extending along these rays,

we have the following:

$$|E_n| \sim \mathcal{K}(n), \tag{26}$$

$$|E_n| \sim \mathcal{K}(n)_{\mathfrak{R},\,\mathfrak{R}+1}. \tag{27}$$

If we consider Fig. 21 (f and h) and especially Fig. 21(j), we will see that the form of the envelope after a large number of passes resembles more of a hyperbole of the following type:

$$E_n \sim -1/n^2. \tag{28}$$

This ratio obtained from our geometric constructions corresponds to experimental results in spectroscopy and theoretical results of Bohr's theory and quantum mechanics [6] and [7].

Fig. 22 shows the results (similar to those that were shown in Fig. 21 f, h, and j) of numerical calculations of layers of the nonlinear arithmetic parallelepiped for the eighth R shell in the form of envelopes of distribution of the ray number. In (a) and (b), you can see the envelopes of distribution of the ray number at the angle $\mathcal{K}(n)$ for passes of rays $\mathfrak{R} = 131$, 132. Results of joint calculations for passes of $\mathfrak{R} = 131$ and $\mathfrak{R} = 132$ in the form of the envelope are given in (c), and these in a form of histograms are given in (d) and (e).

If we consider Fig. 22 (a and b) and especially Fig. 22 (c to e), we will see that the form of the envelope after a large number of passes resembles more of a hyperbole (28).

Fig. 22 (e) similar to Fig. 22 (c and d) should be compared to Fig. 28 of the Appendix.

Our numerical calculations show that with an increase in the number of passes of $\mathfrak{R}$, and an increase in the number of subshells of a shell and the main number n, the form of an envelope, Fig. 21 (j) and Fig. 22 (c), increasingly resembles a hyperbole of (28) type. If the number of subshells exceeds eight (e.g., you can construct eleven), eight of eleven subshells can be subsumed to subshells, while the rest three can be subsumed to a continuous spectrum [6] and [7]. Such creation of a continuous spectrum does not contradict Pauli Principle.

Conclusions

In our illustrative geometric researches, using just one basic summation formula of $A = B + C$ (3), Excel, and various initial and threshold conditions set, we have revealed a number of new regularities like we did in previous works [3] and [4].

It appeared that quantum systems can be geometrically interpreted by means of our model of a half-integer rays system in an illustrative way.

We have described Pauli Principle, shells and subshells of atoms of the periodic table. At the same time, the number of shells and subshells in our model does not exceed eight, and all the subshells starting with the ninth can be considered a continuous spectrum.

By means of our model, it is possible to interpret the principle, azimuthal, magnetic, and spin quantum numbers in the form of angles and distances.

By means of our model, we have given a separate geometric interpretation of an atom of hydrogen and its power levels. We have interpreted transitions of an electron from one level to another in the form of angles, but not distances as it is commonly interpreted [6] and [7]. In this work, we have also shown that the hyperbolic dependence of energy of a hydrogen atom of $E_n \sim -1/n^2$ (28) known from experimental spectral studies, Bohr's theory and quantum mechanics, can be also obtained from our geometric constructions on the basis of Pauli Principle.

Based on the research of a half-integer ray model, we have illustrated the stepped structure of layers at laminar flow of liquid [8]. The similar stepped structure was observed in research of integer ray model made by us in [3], but the half-integer model gives the more accurate image of "steps" in comparison with our integer model.

Acknowledgements

The author expresses gratitude to the late Prof. E. E. Shnoll, Prof. J. Peters (Canada), Prof. V. V. Dikusar (Russia), Prof. S. E. Shnoll (Russia), Prof. A. A. Rukhadze (Russia), Prof. V. G. Mikhalevich (Russia), Prof. A. V. Kaganov (Russia), Prof. R. Mehta (India), Dr. A. Shemetov (Russia), Dr. O. Nersesyan (Russia) and Captain V. Kabanov (Russia) for useful discussion and support. The author is also grateful to the Prof. N. J. A. Sloane (USA) for his surprising tables and to the Excel software creators for their excellent Excel.

Submitted on December 26, 2015 / Accepted on January 19, 2015

References

1. Peters J. F., Tozzi A. The Borsuk-Ulam theorem explanes quantum entangement. *Technical Reports*, November 2015, DOI: 10.13140/RG.2.1.3860.1685.
2. Yurkin A. V. Quasi-resonator a new interpretation of scattering in lasers. *Quantum Electronics*, 1994, v. 24, 359.
3. Yurkin A. V. Symmetric Triangle of Pascal and Arithmetic Parallelepiped. On Possibility of New Evident Geometrical Interpretation of Processes in Long Pipes. Lambert Academic Publishing, 2015.
4. Yurkin A. V. New Binomial and New View on Light Theory. About One New Universal Descriptive Geometric Model. Lambert Academic Publishing, 2013.
5. Yurkin A. V. Recurrence calculation of laser divergence and refractive analog of a multilobe mirror. *Quantum Electronics*, 1993, v. 23, 323.
6. Putilov K. A., Fabrikant V. A. The Course of Physics. Vol. 3. Moscow, Fizmatgiz, 1960 (in Russian).
7. Savelyev I. V. The General Course of Physics. Vol. 3. Moscow, Nauka, 1982 (in Russian).
8. Putilov K. A. The Course of Physics. Vol. 1. Moscow, Fizmatgiz, 1954 (in Russian).

Appendix: reference tables

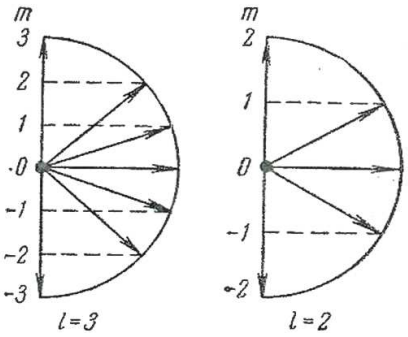

Fig. 23: Illustration of the principle of spatial quantization. Possible values of projections of the orbital momentum to the direction of a magnetic field for $l = 3$ and $l = 2$. (Fig. 231 from [6]).

Shell	n	l	m_l	m_s	Subshell	Shell	n	l	m_l	m_s	Subshell
K	1	0	0	↑↓	$K(1s)$	N	4	0	0	↑↓	$N_1(4s)$
L	2	0	0	↑↓	$L_1(2s)$			1	−1	↑↓	$N_2(4p)$
		1	−1	↑↓	$L_2(2p)$				0	↑↓	
			0	↑↓					+1	↑↓	
			+1	↑↓				2	−2	↑↓	$N_3(4d)$
M	3	0	0	↑↓	$M_1(3s)$				−1	↑↓	
		1	−1	↑↓	$M_2(3p)$				0	↑↓	
			0	↑↓					+1	↑↓	
			+1	↑↓					+2	↑↓	
		2	−2	↑↓	$M_3(3d)$			3	−3	↑↓	$N_4(4f)$
			−1	↑↓					−2	↑↓	
			0	↑↓					−1	↑↓	
			+1	↑↓					0	↑↓	
			+2	↑↓					+1	↑↓	
									+2	↑↓	
									+3	↑↓	

Fig. 24: Division of possible conditions of an electron in an atom into shells and subshells. (Table 36.1 from [7]).

Element	K	L		M			N		Basic therm
	1s	2s	2p	3s	3p	3d	4s	4p	
1 H	1	—	—	—	—	—	—	—	$^2S_{1/2}$
2 He	2	—	—	—	—	—	—	—	$^1S_0$
3 Li	2	1	—	—	—	—	—	—	$^2S_{1/2}$
4 Be	2	2	—	—	—	—	—	—	$^1S_0$
5 B	2	2	1	—	—	—	—	—	$^2P_{1/2}$
6 C	2	2	2	—	—	—	—	—	$^3P_0$
7 N	2	2	3	—	—	—	—	—	$^4S_{3/2}$
8 O	2	2	4	—	—	—	—	—	$^3P_2$
9 F	2	2	5	—	—	—	—	—	$^2P_{3/2}$
10 Ne	2	2	6	—	—	—	—	—	$^1S_0$
11 Na	2	8		1	—	—	—	—	$^2S_{1/2}$
12 Mg	2	8		2	—	—	—	—	$^1S_0$
13 Al	2	8		2	1	—	—	—	$^2P_{1/2}$
14 Si	2	8		2	2	—	—	—	$^3P_0$
15 P	2	8		2	3	—	—	—	$^4S_{3/2}$
16 S	2	8		2	4	—	—	—	$^3P_2$
17 Cl	2	8		2	5	—	—	—	$^2P_{3/2}$
18 Ar	2	8		2	6	—	—	—	$^1S_0$
19 K	2	8		8		—	1	—	$^2S_{1/2}$
20 Ca	2	8		8		—	2	—	$^1S_0$
21 Sc	2	8		8		1	2	—	$^2D_{3/2}$
22 Ti	2	8		8		2	2	—	$^3F_2$
23 V	2	8		8		3	2	—	$^4F_{3/2}$
24 Cr	2	8		8		5	1	—	$^7S_3$
25 Mn	2	8		8		5	2	—	$^6S_{5/2}$
26 Fe	2	8		8		6	2	—	$^5D_4$
27 Co	2	8		8		7	2	—	$^4F_{9/2}$
28 Ni	2	8		8		8	2	—	$^3F_4$
29 Cu	2	8		8		10	1	—	$^2S_{1/2}$
30 Zn	2	8		8		10	2	—	$^1S_0$
31 Ga	2	8		8		10	2	1	$^2P_{1/2}$
32 Ge	2	8		8		10	2	2	$^3P_0$
33 As	2	8		8		10	2	3	$^4S_{3/2}$
34 Se	2	8		8		10	2	4	$^3P_2$
35 Br	2	8		8		10	2	5	$^2P_{3/2}$
36 Kr	2	8		8		10	2	6	$^1S_0$

Fig. 25: The process of building electron shells of the first 36 elements of the periodic system. (Table 37.1 from [7]).

Appendix: reference tables (continue)

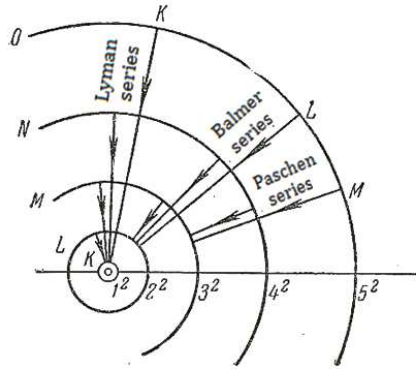

Fig. 26: Orbits of a hydrogen atom in Bohr's theory. The radial arrows located between circles show transitions of an electron from one level to another. (Fig. 228 from [6]).

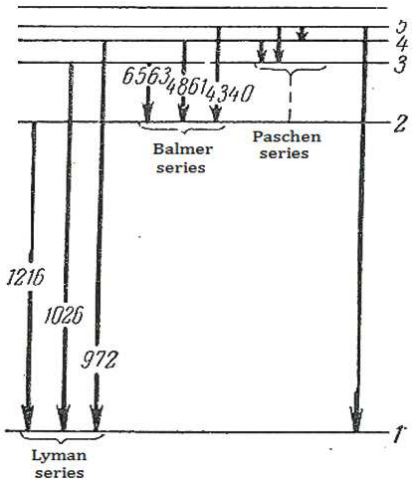

Fig. 27: Scheme of levels of energy of a hydrogen atom. The vertical arrows located between horizontal lines show transitions of an electron from one level to another. (Fig. 229 from [6]).

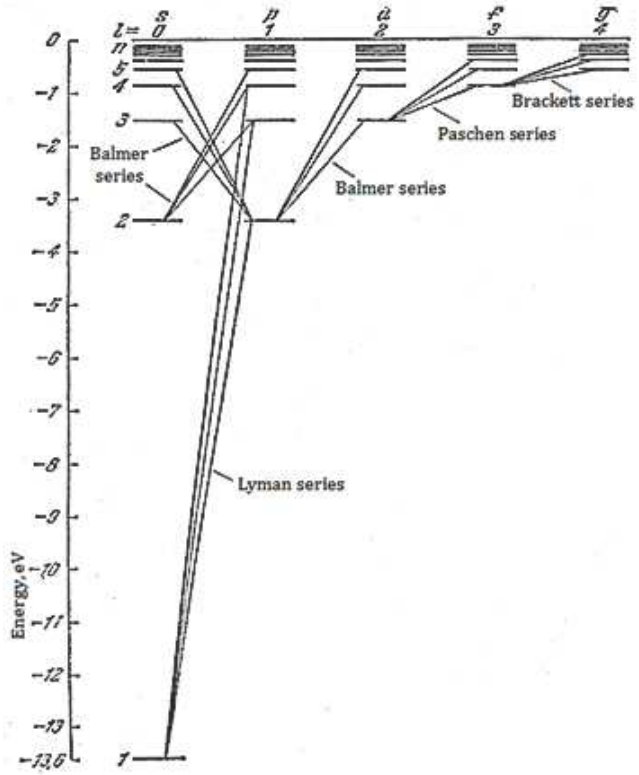

Fig. 28: Scheme of levels of energy of a hydrogen atom. The inclined lines located between horizontal lines show transitions of an electron from one level to another according to the rule of selection $\Delta l = \pm 1$. It means that only transitions upon which l changes by unit are possible. (Fig. 28.1 from [7]).

$X(5)$ Symmetry to ^{152}Sm

Salah A. Eid[1] and Sohair M. Diab[2]

[1]Faculty of Engineering, Phys. Dept., Ain Shams University, Cairo, Egypt.
[2]Faculty of Education, Phys. Dept., Ain Shams University, Cairo, Egypt.
E-mail: mppe2@yahoo.co.uk

The excited positive and negative parity states, potential energy surfaces, $V(\beta,\gamma)$, electromagnetic transition probabilities, $B(E1)$, $B(E2)$, electric monopole strength $X(E0/E2)$ and staggering effect, $\Delta I = 1$, were calculated successfully using the interacting boson approximation model IBA-1. The calculated values are compared to the available experimental data and show reasonable agreement. The energy ratios and contour plot of the potential energy surfaces show that ^{152}Sm is an $X(5)$ candidate.

1 Introduction

Phase transition is one of the very interesting topic in nuclear structure physics. The even-even samarium series of isotopes have encouraged many authors to study that area extensively experimentally and theoretically.

Experimentally, authors studied levels energy with their half-lives, transition probabilities, decay schemes, multipole mixing ratios, internal conversion coefficients, angular correlations and nuclear orientation of γ-rays[1-4].

Theoretically, different theoretical models have been applied to that chain of isotopes. One of the very interesting models is the interacting boson approximation model IBA [5-10]. Iachello [11,12] has made an important contribution by introducing the new dynamical symmetries $E(5)$ and $X(5)$.

$E(5)$ is the critical point symmetry of phase transition between $U(5)$ and $O(6)$ while $X(5)$ is between $U(5)$ and $SU(3)$ nuclei. The aim of the present work is to calculate:

1. The potential energy surfaces, $V(\beta,\gamma)$;
2. The levels energy, electromagnetic transition rates $B(E1)$ and $B(E2)$;
3. The staggering effect, and
4. The electric monopole strength $X(E0/E2)$.

2 IBA-1 model

2.1 Levels energy

The IBA-1 Hamiltonian [13-16] employed on ^{152}Sm in the present calculation is:

$$\begin{aligned} H &= EPS\cdot n_d + PAIR\cdot(P\cdot P) \\ &+\frac{1}{2}\,ELL\cdot(L\cdot L)+\frac{1}{2}\,QQ\cdot(Q\cdot Q) \\ &+5\,OCT\cdot(T_3\cdot T_3)+5HEX\cdot(T_4\cdot T_4)\,, \end{aligned} \tag{1}$$

where

$$P\cdot p=\frac{1}{2}\begin{bmatrix} \left\{(s^\dagger s^\dagger)^{(0)}_0-\sqrt{5}(d^\dagger d^\dagger)^{(0)}_0\right\}x \\ \left\{(ss)^{(0)}_0-\sqrt{5}(\tilde{d}\tilde{d})^{(0)}_0\right\}\end{bmatrix}^{(0)}_0, \tag{2}$$

$$L\cdot L=-10\sqrt{3}\left[(d^\dagger\tilde{d})^{(1)}x\,(d^\dagger\tilde{d})^{(1)}\right]^{(0)}_0\,, \tag{3}$$

$$Q\cdot Q=\sqrt{5}\begin{bmatrix} \left\{(S^\dagger\tilde{d}+d^\dagger s)^{(2)}-\dfrac{\sqrt{7}}{2}\,(d^\dagger\tilde{d})^{(2)}\right\}x \\ \left\{(s^\dagger\tilde{d}++\tilde{d}s)^{(2)}-\dfrac{\sqrt{7}}{2}\,(d^\dagger\tilde{d})^{(2)}\right\}\end{bmatrix}^{(0)}_0, \tag{4}$$

$$T_3\cdot T_3=-\sqrt{7}\left[(d^\dagger\tilde{d})^{(2)}x\,(d^\dagger\tilde{d})^{(2)}\right]^{(0)}_0\,, \tag{5}$$

$$T_4\cdot T_4=3\left[(d^\dagger\tilde{d})^{(4)}x\,(d^\dagger\tilde{d})^{(4)}\right]^{(0)}_0\,. \tag{6}$$

In the previous formulas, n_d is the number of bosons; $P\cdot P$, $L\cdot L$, $Q\cdot Q$, $T_3\cdot T_3$ and $T_4\cdot T_4$ represent pairing, angular momentum, quadrupole, octupole and hexadecupole interactions respectively between the bosons; EPS is the boson energy; and $PAIR$, ELL, QQ, OCT, HEX are the strengths of the pairing, angular momentum, quadrupole, octupole and hexadecupole interactions respectively (see Table 1).

2.2 Transition rates

The electric quadrupole transition operator employed is:

$$\begin{aligned} T^{(E2)} &= E2S\,D\cdot(s^\dagger\tilde{d}+d^\dagger s)^{(2)}+ \\ &+\frac{1}{\sqrt{5}}\,E2DD\cdot(d^\dagger\tilde{d})^{(2)}\,. \end{aligned} \tag{7}$$

$E2S\,D$ and $E2DD$ are adjustable parameters.

The reduced electric quadrupole transition rates between $I_i\to I_f$ states are given by:

$$B(E_2,I_i-I_f)=\frac{[<I_f\parallel T^{(E_2)}\parallel I_i>]^2}{2I_i+1}\,. \tag{8}$$

3 Results and discussion

In this section we review and discuss the results.

nucleus	*EPS*	*PAIR*	*ELL*	*QQ*	*OCT*	*HEX*	*E2S D*(*eb*)	*E2DD*(*eb*)
^{152}Sm	0.3840	0.000	0.0084	−0.0244	0.0000	0.0000	0.1450	−0.4289

Table 1: Parameters used in *IBA*-1 Hamiltonian (all in MeV).

3.1 The potential energy surfaces

The potential energy surfaces [17], $V(\beta, \gamma)$, as a function of the deformation parameters β and γ are calculated using:

$$\begin{aligned} E_{N_\Pi N_\nu}(\beta,\gamma) &= <N_\pi N_\nu;\beta\gamma|H_{\pi\nu}|N_\pi N_\nu;\beta\gamma> = \\ &= \zeta_d(N_\nu N_\pi)\beta^2(1+\beta^2)+\beta^2(1+\beta^2)^{-2}\times \\ &\times\left\{kN_\nu N_\pi[4-(\bar{X}_\pi\bar{X}_\nu)\beta\cos 3\gamma]\right\}+ \\ &+\left\{[\bar{X}_\pi\bar{X}_\nu\beta^2]+N_\nu(N_\nu-1)\left(\frac{1}{10}\,c_0+\frac{1}{7}\,c_2\right)\beta^2\right\}, \end{aligned} \tag{9}$$

where

$$\bar{X}_\rho=\left(\frac{2}{7}\right)^{0.5} X_\rho\,\rho=\pi \text{ or } \upsilon\,. \tag{10}$$

The calculated potential energy surfaces, $V(\beta,\gamma)$, are presented in Figures 1, 2, 3. ^{152}Sm lies between ^{150}Sm which is a vibrational like nucleus, $U(5)$, Fig. 1, while ^{154}Sm is a rotational like, $SU(3)$, nucleus, Fig. 3. So, ^{150}Sm can be an $X(5)$ candidate where levels energy, transition probability ratios as well as the potential energy surfaces are supporting that assumption (see Table 2).

3.2 Energy spectra and electric transition rates

The energy of the positive and negative parity states of ^{152}Sm isotope are calculated using computer code PHINT [19]. A comparison between the experimental spectra [18] and our calculations, using values of the model parameters given in Table 1 for the ground state, $\beta1$, $\beta2$ and γ bands are illustrated in Fig. 4. The agreement between the calculated levels energy and their corresponding experimental values are fair, but they are slightly higher especially for the higher excited states in $\beta1$, $\beta2$ and γ bands. We believe this is due to the change of the projection of the angular momentum which is due mainly to band crossing. Fig. 5 shows the position of $X(5)$ and $E(5)$ between the other types of nuclei.

Unfortunately there are no available measurements of electromagnetic transition rates $B(E1)$ for ^{152}Sm nucleus, Table 3, while some of $B(E2)$ are measured. The measured $B(E2, 2_1^+ \to 0_1^+)$ is presented, in Table 4, for comparison with the calculated values [20]. The parameters *E2S D* and *E2DD* displayed in Table 1 are used in the computer code NPBEM [19] for calculating the electromagnetic transition rates and the calculated values are normalized to $B(E2, 2_1^+ \to 0_1^+)$. No new parameters are introduced for calculating electromagnetic transition rates $B(E1)$ and $B(E2)$ of intraband and interband.

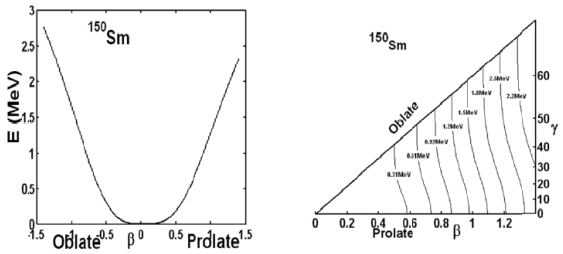

Fig. 1: Potential energy surfaces for ^{150}Sm .

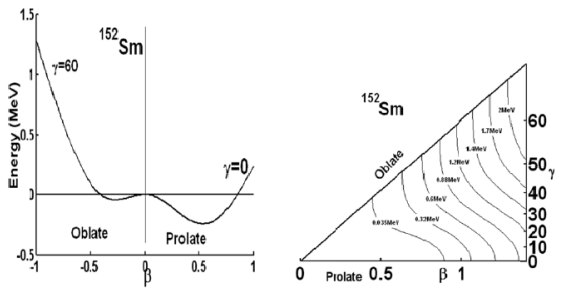

Fig. 2: Potential energy surfaces for ^{152}Sm .

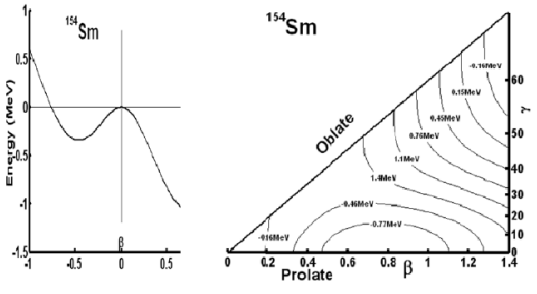

Fig. 3: Potential energy surfaces for ^{154}Sm .

nucleus	$E_{4_1^+}/E_{2_1^+}$	$E_{6_1^+}/E_{2_1^+}$	$E_{8_1^+}/E_{2_1^+}$	$E_{0_2^+}/E_{2_1^+}$	$E_{6_1^+}/E_{0_2^+}$	$E_{0_3^+}/E_{2_1^+}$	$BE2(4_1^+-2_1^+)/BE2(2_1^+-0_1^+)$
^{152}Sm	3.02	5.83	9.29	5.66	1.03	8.92	1.53
$X(5)$	3.02	5.83	9.29	5.65	1.53	6.03	1.58

Table 2: Energy and transition probability ratios.

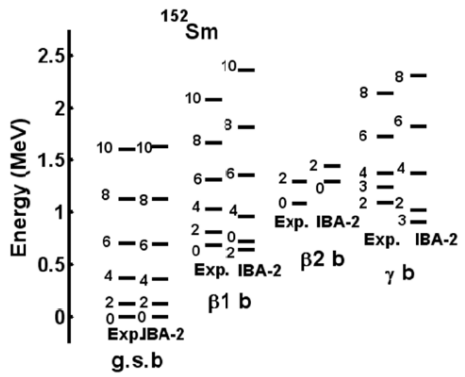

Fig. 4: Experimental[18] and calculated levels energy

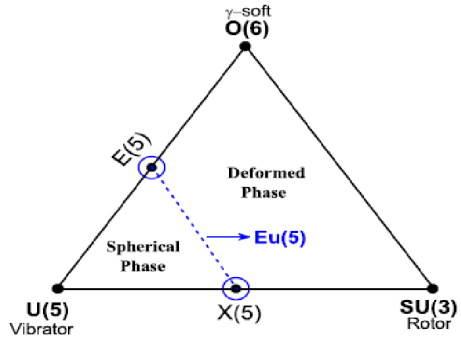

Fig. 5: Triangle showing the position of $X(5)$ and $E(5)$.

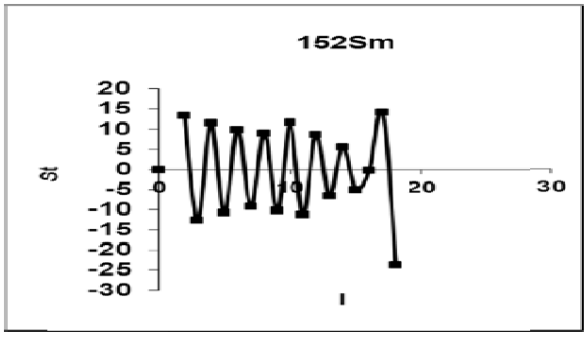

Fig. 6: Staggering effect on ^{152}Sm.

I_i^- I_f^+	$B(E1)$Exp.	$B(E1)$IBA-1
1_1 0_1	—-	0.0979
1_1 0_2	—-	0.0814
3_1 2_1	—-	0.2338
3_1 2_2	—-	0.0766
3_1 2_3	—-	0.0106
3_2 2_1	—-	0.0269
3_2 2_2	—-	0.0291
3_2 2_3	—-	0.0434
5_1 4_1	—-	0.3579
5_1 4_2	—-	0.0672
5_1 4_3	—-	0.0050
7_1 6_1	—-	0.4815
7_1 6_2	—-	0.0574
9_1 8_1	—-	0.6075
9_1 8_2	—-	0.0490
11_1 10_1	—-	0.7367
11_1 10_2	—-	0.0413

Table 3: Calculated $B(E1)$ in ^{152}Sm.

3.3 Staggering effect

The presence of $(+ve)$ and $(-ve)$ parity states has encouraged us to study the staggering effect [21-23] for ^{152}Sm isotope using staggering function equations (11, 12) with the help of the available experimental data [18].

$$St(I) = 6\Delta E(I) - 4\Delta E(I-1) - 4\Delta E(I+1) + \\ + \Delta E(I+2) + \Delta E(I-2)\,, \qquad (11)$$

with

$$\Delta E(I) = E(I+1) - E(I)\,. \qquad (12)$$

The calculated staggering patterns are illustrated in Fig. 6 and show an interaction between the $(+ve)$ and $(-ve)$ parity states for the ground state band of ^{152}Sm.

3.4 Electric monopole transitions

The electric monopole transitions, $E0$, are normally occurring between two states of the same spin and parity by transferring energy and zero unit of angular momentum. The strength of the electric monopole transition, $X_{if'f}(E0/E2)$, [24] can be calculated using equations (13, 14) and presented in Table 5.

$$X_{if'f}(E0/E2) = \frac{B(E0, I_i - I_f)}{B(E2, I_i - I_{'f})}\,, \qquad (13)$$

I_i^+ I_f^+	$B(E2)$Exp*.	$B(E2)$IBA-1
2_1 0_1	0.670(15)	0.6529
3_1 2_1	—-	0.0168
4_1 2_1	0.1.017(4)	1.0014
6_1 4_1	1.179(33)	1.1304
0_2 2_1	0.176(1)	0.3363
2_2 2_1	0.0258(26)	0.0610
2_2 4_1	0.091(11)	0.1057
4_2 2_1	0.0035(35)	0.0003
4_2 4_1	0.037(23)	0.0458
2_3 0_1	0.0163(11)	0.0141
2_3 2_1	0.0417(42)	0.0125
2_3 4_1	0.0416(32)	0.0296
4_3 2_1	0.0035(13)	0.0038
4_3 4_1	0.037(13)	0.0084
4_3 4_2	—-	0.1235
4_3 2_2	—-	0.0070
4_3 2_3	—-	0.3110
4_2 2_2	—-	0.6418
8_1 6_1	—-	1.1681
8_1 6_2	—-	0.0376
10_1 8_1	—-	1.1421

Table 4: Calculated $B(E2)$ in ^{152}Sm (* from Ref.[20])

I_i^+ I_f^+	X(E0/E2)Exp*.	X(E0/E2)IBA-1
0_2 0_1	0.7(0.1)	0.85
0_3 0_2	—-	3.68
0_3 0_1	—-	0.72
0_4 0_3	—-	4.39
0_4 0_2	—-	0.64
0_4 0_1	—-	1.27
2_2 2_1	4.5(0.5)	3.52
2_3 2_1	—-	12.23
2_3 2_2	—-	11.19
4_3 4_1	—-	1.76
4_3 4_2	—-	1.40
4_4 4_1	—-	0.44
4_4 4_2	—-	3.15
4_2 4_1	6.6(2.10)	2.02
6_2 6_1	—-	1.46
8_2 8_1	—-	1.20
10_2 10_1	—-	1.07

Table 5: $X_{if'f}$ $(E0/E2)$ ratios in ^{152}Sm (* from Ref [20]).

where I_i =I_f=0 , $I_{'f}$=2 and I_i= $I_f \neq 0$, I_f = $I_{'f}$.

$$X_{if'f}(E0/E2) = (2.54{\times}10^9)\,A^{3/4} \times \times \frac{E_\gamma^5(\mathrm{MeV})}{\Omega_{KL}}\,\alpha(E2)\,\frac{T_e(E0, I_i - I_f)}{T_e(E_2, I_i - I_{'f})}\,. \qquad (14)$$

where:
A : mass number;
I_i : spin of the initial state where E0 and E2 transitions are depopulating it;
I_f : spin of the final state of E0 transition;
$I_{'f}$: spin of the final state of E2 transition;
E_γ : gamma ray energy;
Ω_{KL} : electronic factor for K,L shells [25];
$\alpha(E2)$: conversion coefficient of the E2 transition;
$T_e(E0, I_i - I_f)$: absolute transition probability of the $E0$ transition between I_i and I_f states, and
$T_e(E_2, I_i - I_{'f})$: absolute transition probability of the $E2$ transition between I_i and $I_{'f}$ states.

3.5 Conclusions

The IBA-1 model has been applied successfully to the ^{152}Sm isotope and:

1. Levels energy are successfully reproduced;
2. Potential energy surfaces are calculated and show $X(5)$ characters to ^{152}Sm;
3. Electromagnetic transition rates $B(E1)$ and $B(E2)$ are calculated;
4. Staggering effect has been calculated and beat pattern observed which show an interaction between the $(-ve)$ and $(+ve)$ parity states, and
5. Strength of the electric monopole transitions $X_{if'f}(E0/E2)$ are calculated.

Submitted on January 10, 2016 / Accepted on January 12, 2016

References

1. Qian Y., Ren Z. and Ni D. α-decay half-lives in medium mass nuclei. *J. Phys. G*, 2010, v. 38, 1.
2. Palla G., Geramb H. V. and Pegel C. Electric and inelastic scattering of 25.6 MeV protons from even samarium isotopes. *Nucl. Phys. A*, 1983, v. 403, 134.
3. Talon P., Alamanos N., Lamehi-Rachti M., Levi C. and Papineau L. Coulomb and nuclear excitation effects on ^{16}O scattering from samarium isotopes. *Nucl. Phys. A*, 1981, v. 359, 493.
4. Palla G. and Pegei C. Inelastic scattering of hellions from even-even stable samarium isotopes at 40.9 MeV. *Nucl. Phys. A*, 1979, v. 321, 317.
5. Sabri H. Spectral statistics of rare-earth nuclei: Investigation of shell model configuration effect. *Nucl. Phys. A*, 2015, v. 941, 364.
6. Nomura K., Vretenar D., Nikie T. and Lu B. Microscopic description of octupole shape-phase transitions in light actinide and rare-earth nuclei. *Phys. Rev. C*, 2014, v. 89, 024312.
7. Nikie T., Kralj N., Tutis T., Vretenar D., and Ring P. Implementation of the finite amplitude method for the relativistic quasiparticle random-phase approximation *Phys. Rev. C*, 2013, v. 88, 044327.

8. Nomura K. Interacting boson model with energy density functionals. *J. Phys. Conference Series*, 2013, v. 445, 012015.
9. Nomura K., Shimizu N., and Otsuka T. Formulating the interacting boson model by mean-field methods. *Phys. Rev. C*, 2010, v. 81, 044307.
10. Zhang W., Li Z. P., Zhang S. Q., and Meng J. Octupole degree of freedom for the critical-point candidate nucleus Sm152 in a reflection-asymmetric relativistic mean-field approach. *Phys. Rev. C*, 2010, v. 81, 034302.
11. Iachello F. Dynamic symmetries of the critical point. *Phys. Rev. Lett.*, 2000, v. 85, 3580.
12. Iachello F. Analytic describtion of critical point nuclei in spherical-axially deformed shape phase transition. *Phys. Rev. Lett.*, 2000, v. 85, 3580.
13. Puddu G., Scholten O., Otsuka T. Collective Quadrupole States of Xe, Ba and Ce in the Interacting Boson Model. *Nucl. Phys.*, 1980, v. A348, 109.
14. Arima A. and Iachello F. Interacting boson model of collective states: The vibrational limit. *Ann. Phys.*, 1976, v. 99, 253.
15. Arima A. and Iachello F. Interacting boson model of collective states: The rotational limit. *Ann. Phys.*, 1978, v. 111, 201.
16. Arima A. and Iachello F. Interacting boson model of collective states: The $O(6)$ limit. *Ann. Phys.*, 1979, v. 123, 468.
17. Ginocchio J. N. and Kirson M. W. An intrinisic state for the interacting boson model and its relationship to the Bohr-Mottelson approximation. *Nucl. Phys. A*, 1980, v. 350, 31.
18. Martin M.J. Adopted levels, gammas for ^{152}Sm. *Nucl. Data Sheets*, 2013, v. 114, 1497.
19. Scholten O. The programm package PHINT (1980) version, internal report KVI-63. Keryfysisch Versneller Institut, Gronigen, 1979.
20. Church E. L. and Weneser J. Electron monopole transitions in atomic nuclei. *Phys. Rev.*, 1956, v. 103, 1035.
21. Minkov N., Yotov P., Drenska S. and Scheid W. Parity shift and beat staggering structure of octupole bands in a collective model for quadrupole-octupole deformed nuclei. *J. Phys. G*, 2006, v. 32, 497.
22. Bonatsos D., Daskaloyannis C., Drenska S. B., Karoussos N., Minkov N., Raychev P. P. and Roussev R. P. $\Delta I = 1$ staggering in octupole bands of light actinides Beat patterns. *Phys. Rev. C*, 2000, v. 62, 024301.
23. Minkov N., Drenska S. B., Raychev P. P., Roussev R. P. and Bonatsos D. Beat patterns for the odd-even staggering in octupole bands from quadrupole-octupole Hamiltonian. *Phys. Rev. C*, 2001, v. 63, 044305.
24. Rasmussen J. O. Theory of $E0$ transitions of spheroidal nuclei. *Nucl.Phys.*, 1960, v. 19, 85.
25. Bell A. D., Aveldo C. E., Davidson M. G. and Davidson J. P. Table of $E0$ conversion probability electronic factors. *Can. J. Phys.*, 1970, v. 48, 2542.

LETTERS TO PROGRESS IN PHYSICS

A Re-examination of Kirchhoff's Law of Thermal Radiation in Relation to Recent Criticisms

Robert J. Johnson
E-mail: bob.johnson1000@gmail.com

This paper investigates claims made by Pierre-Marie Robitaille in a series of papers from 2003 to 2015 that Kirchhoff's Law of thermal radiation does not apply to cavities made of arbitrary materials, and that Planck's theoretical derivation and apparent proof of this law in these cases is faulty. Robitaille's claims are compared to statements in the original papers by Kirchhoff and Planck. The present paper concludes that Robitaille's claims are not sustainable and that Kirchhoff's Law and Planck's proof remain valid in the situations for which they were intended to apply, including in cavities with walls of any arbitrary materials in thermal equilibrium.

1 Introduction

In a series of papers from 2003 to 2015 [1–10], Pierre-Maire Robitaille has challenged the validity of Kirchhoff's Law of thermal emission and Planck's derivation of the mathematical form of the universal function of spectral radiance absorbed and emitted by a black body. As the consequences of a failure of Kirchhoff's Law would, if proven, include the loss of universality of application of various fundamental physical constants including *'Planck's constant, Boltzmann's constant, ... "Planck length", "Planck time", "Planck mass", and "Planck temperature"'* [10, p. 121], Robitaille's claims deserve serious consideration.

In this paper, Robitaille's claims will be compared to the original works by Kirchhoff [11] and Planck [12] in order to determine whether his criticisms of these earlier works are valid. The present paper focuses initially on the arguments contained in the series of papers by Robitaille from 2003 to 2014 [1–9]; the second part will address the recent paper authored jointly by Robitaille and Crothers [10].

2 Robitaille's earlier papers [1–9]

2.1 Kirchhoff's law and Planck's proof

Kirchhoff's Law of thermal radiation dating from 1859-1860 may be stated as follows: *"For an arbitrary body radiating and emitting thermal radiation, the ratio E / A between the emissive spectral radiance, E, and the dimensionless absorptive ratio, A, is one and the same for all bodies at a given temperature. That ratio E / A is equal to the emissive spectral radiance I of a perfect black body, a universal function only of wavelength and temperature"*. This radiance, I, is often referred to simply as black radiation.

The form of the universal function was not known until Planck derived it theoretically in 1914 in what is now known as Planck's Law. Planck's derivation is seen as proof of Kirchhoff's Law. However, Robitaille points out that the above definition of Kirchhoff's Law is not complete and furthermore Robitaille maintains that the statement above should be called Stewart's Law as it was originally propounded by Stewart in 1858 [13]: *"All too frequently, the simple equivalence between apparent spectral absorbance and emission is viewed as a full statement of Kirchhoff's law, ... Kirchhoff's law must always be regarded as extending much beyond this equivalence. It states that the radiation within all true cavities made from arbitrary walls is black. The law of equivalence is Stewart's"* [5, p. 11].

According to Robitaille, in deriving his law of equivalence Stewart had considered the case of a cavity made from perfectly absorbing (i.e. black) material; he had shown that the radiation in such a cavity at thermal equilibrium must also be black, of an intensity appropriate to the equilibrium temperature.

Whilst Robitaille agrees with Stewart, he profoundly disagrees with Kirchhoff's extension of this finding to cavities made of arbitrary materials, and therefore with Planck's proof of Kirchhoff's result. Planck had based his proof on a consideration of perfectly reflecting cavities containing *"an arbitrarily small quantity of matter"* [12, § 51], arriving at the same result that Kirchhoff had obtained for perfectly absorbing cavities. Planck had thereby demonstrated that all cavities either containing some arbitrary matter, or equivalently having walls made of some arbitrary matter, must also contain black radiation when at thermal equilibrium.

2.2 Black radiation in a perfectly reflecting cavity

In following the reasoning of both sides of this disagreement, it is important to distinguish between a perfectly reflecting cavity containing a vacuum and one containing an opaque object or a partially-absorbing medium.

In the first case, Planck writes that *"Hence in a vacuum*

bounded by totally reflecting walls any state of radiation may persist" [12, § 51]; Robitaille claims that this statement is a violation of Kirchhoff's Law [10, p. 130]. However, Planck's statement should perhaps be more properly be viewed as a situation to which Kirchhoff's Law does not apply because there is no matter present which either absorbs or emits radiation.

When considering the case of a perfectly reflecting cavity containing an arbitrary object, again it is important to distinguish between two situations. The first is that the object absorbs and emits some fraction of all frequencies of radiation; this situation may be further subdivided into the special case where the object is itself a black body such as Planck's particle of carbon which is a perfect absorber and emitter at all frequencies; and the general case where the object only absorbs and emits some fraction above zero but less than unity of every frequency. The second situation is that the object only absorbs and emits over part of the spectrum i.e. there are some frequencies for which the object itself is a perfect reflector, neither absorbing nor emitting at those frequencies.

The question in both situations is, what is the nature of the radiation in the perfectly reflecting cavity at thermal equilibrium?

Starting with the special case of a black body, Robitaille, Kirchhoff and Planck all agree that the radiation is necessarily black. The disagreements start over the general case of an object imperfectly absorbing at all frequencies.

Planck maintains that *"... the radiation of a medium completely enclosed by absolutely reflecting walls is, when thermodynamic equilibrium has been established for all colors for which the medium has a finite coefficient of absorption, always the stable radiation corresponding to the temperature of the medium such as is represented by the emission of a black body"* [12, § 51], quoted in [1, p. 1263]. Note that the quoted statement covers both the situation where the object absorbs and emits over all frequencies, and the situation where some frequencies are not absorbed or emitted at all.

In contrast, Robitaille claims that *"In fact, if an object is placed within* [perfectly reflecting] *walls, an equilibrium will be established, but it will not correspond to that of a blackbody. Indeed, the radiation contained within such a device will reflect purely the emission profile of the object it contains"* [1, p. 1264].

This is Robitaille's central argument against the universality claimed by Kirchhoff and Planck i.e. that all cavities containing an object must, at equilibrium, come to contain black radiation at all frequencies absorbed and emitted by the object.

2.3 The approach to equilibrium

In effect, the argument comes down to the quantity of the radiation in the cavity at equilibrium. Both sides agree that there is some radiation at all frequencies absorbed and emitted by the object; the disagreement is over the intensity of that radiation. Does it, as Kirchhoff and Planck maintain, equal the intensity of black radiation which we can now quantify according to Planck's Law of 1914; or does the radiation density in the cavity fall short of the black body level at some or all frequencies because of the imperfect absorption and emission of the object in the cavity, as Robitaille claims?

The role played by the reflected radiation, i.e. that fraction of incident radiation which is not fully absorbed by the object, is the key. Robitaille maintains that the radiation density in the cavity cannot be increased to black body levels by what he terms "driving the reflection" because this would imply a departure from thermal equilibrium which, Robitaille argues, contravenes the initial assumption that thermal equilibrium exists.

A simplified numerical example may be helpful here in order to crystallise the arguments. Suppose an opaque object in a perfectly reflecting cavity is in thermal equilibrium at a certain temperature and has a coefficient of absorption of 0.8 (i.e. 80%) of all incident radiation at all frequencies. The remaining 20% of any incident radiation will be reflected. Suppose further that the radiation density in the cavity is already at the level at which a black body at the same temperature would be in thermal equilibrium with it, say100 units. This will represent the incident radiation on the opaque object which will then absorb 80 units and reflect 20 units. The object will also re-emit the same 80 units into the cavity. The total radiation coming off the surface of the object, consisting of the emitted and reflected components, is 100, therefore thermal equilibrium will be maintained with the radiation in the cavity. What's more, the radiation density is and remains black according to the initial assumption. This represents the situation described by Kirchhoff's Law.

Consider now the situation where the same object at the same initial temperature is introduced into the perfectly reflecting but otherwise empty cavity, i.e. there is no radiation density in the cavity initially. In this case, the object will emit 80 units appropriate to its temperature; these will be reflected off the walls and become "incident" radiation on the object. The object will now absorb 80%, or 64 units, and reflect 16 units. But it is bound by its initial temperature to continue emitting 80 units. There is therefore a shortfall between the amount absorbed and the amount emitted and the object will cool down. The energy lost by the object will be converted to additional radiation density in the cavity which will increase until equilibrium is achieved between the object and the radiation density at some new, lower, temperature. At this point, the radiation will again be black, but at the level appropriate to the lower temperature, not the initial temperature of the object.

Robitaille would object to this second example on the grounds that thermal equilibrium has not been maintained. This is correct. But Robitaille goes further and maintains that this proves that the cavity cannot contain black radiation because it is not allowable to "drive the reflection" until

a new equilibrium is reached – the object must be maintained at the original temperature throughout and therefore there is no spare energy available to "drive the reflection" up to black body densities.

In essence, Robitaille disallows the approach to thermal equilibrium between the object and the radiation density in the cavity by the mechanism outlined in the second numerical example above. As a result, Robitaille maintains that the cavity cannot contain the black radiation required by Kirchhoff's Law and therefore the law fails.

In support of his argument, Robitaille quotes Stewart [13] as follows: *"Let us suppose we have an enclosure whose walls are of any shape, or any variety of substances (all at a uniform temperature), the normal or statical condition will be, that the heat radiated and reflected together, which leaves any portion of the surface, shall be equal to the radiated heat which would have left that same portion of the surface, if it had been composed of lampblack ... Let us suppose, for instance, that the walls of this enclosure were of polished metal then only a very small quantity of heat would be radiated; but this heat would be bandied backwards and forwards between the surfaces, until the total amount of radiated and reflected heat together became equal to the radiation of lampblack"* [13, § 32] quoted in [4, p. 45].

Robitaille comments: *"These passages are quite similar to Kirchhoff's with the distinction that universality is never invoked. Stewart realizes that the lampblack surface within the enclosure is essential"* [4, p. 45]. But Stewart is quite specific – the walls may be of any variety of substance including polished metal. This implies that Robitaille's objection to what he refers to as Kirchhoff's extension of Stewart's result to cavities made of arbitrary material is unfounded; Stewart had already made the theoretical leap.

How then did Stewart conclude as he did that *"the sum of the radiated and reflected heat together became equal to the radiation of lampblack?"*

2.4 Stewart's treatment of reflection

In Stewart's original paper there is a footnote to the section quoted above which explains the calculation by which he arrived at this conclusion. Stewart considers *"two parallel plates of polished metal of the same description radiating to one another"* [13, § 32-footnote] and investigates what happens to an initial amount r of radiation emitted by each opposing plate and falling perpendicularly on the other plate, where a proportion is reflected back to the first plate. As an ever-decreasing part of the original radiation r is "bandied about" by repeated reflection between the plates, with a proportion $\alpha(< 1)^*$ of the incident radiation being reflected each time, Stewart shows that the total amount falling on one of the plates is

$$r\left(1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4 \ldots\right) = \frac{r}{1-\alpha},$$

which, Stewart explains, is the same formula as results from the case where one of the plates is a black body in thermal equilibrium with the other plate.

The question then arises, can this calculation also be applied to a situation where thermal equilibrium has not yet been achieved? It turns out that it can. Note that, in modern parlance, Stewart's calculation sums the repeated reflections of the two initial pulses (one from each plate) emitted in the first interval of time δt over subsequent intervals of time. It may be supposed without loss of generality that the interval of time δt corresponds to the transmission time of radiation between the plates. Then the same sum would result from considering what proportion of a series of identical initial pulses each of emission duration δt fell on one plate in a single (later) interval of time δt. This second case represents continuous emission of radiation in thermal equilibrium. One of the plates may then be replaced with a black body at the same equilibrium temperature which emits exactly the same amount of radiation that it absorbs, or alternatively with a perfect reflector. Again, the same sum emerges from the calculation, as Stewart explained.

What's more, exactly the same result is obtained when one plate is perfectly reflecting and there is no radiation in the gap between the plates initially, i.e. there is no initial thermal equilibrium to supply the series of constant pulses prior to the arrival interval δt under consideration. In this case, all the radiation is emitted by just one of the plates; therefore double the time is required to achieve the same result that Stewart obtained but, in effect, this result shows that once a steady state has been achieved then the radiation arriving on the single partially-absorbing plate is equivalent to that coming from a black body. The only difference in this case is that during the initial period the partially-absorbing plate is absorbing less radiation than it is emitting; it is therefore cooling down and part of its initial thermal energy is being used to increase the radiation density between the plates, or, in Robitaille's terms, in "driving the reflection". However, when thermal equilibrium is established then the calculation shows that the radiation reflected back on to the emitting plate will be equivalent to black radiation at the equilibrium temperature.

What this demonstrates is that Stewart's method of calculation of the reflection being "bandied about" can also be applied to the approach to equilibrium provided that time is allowed for a sufficient number of reflections to build up the radiation density in the cavity to equilibrium levels. The total time necessary to fill the space with black radiation is likely to be short because of the extremely short transmission time δt and the limited number of reflections necessary to achieve near-perfect black body radiation in most normal situations.

*Stewart uses α to represent the proportion of reflected radiation; in Planck's usage, α represents the coefficient of absorption. To comply with Planck's usage, α should be replaced with ρ in the above equation. The derivation of the equation is unaffected.

Only in cases where the plate is nearly a perfect reflector might an appreciable time be required.

Thus "Stewart's mechanism", if we may so call it, should be interpreted as indicated in the second numerical example given above, with the walls themselves taking the part of the opaque object. Stewart's words "bandied about" can be applied to the reflected proportions of the continuing emission which build up the radiation density in the cavity until thermal equilibrium is achieved. Robitaille calls this "driving the reflection"; it may be clearer to think of the effect as "increasing the radiation energy density in the cavity" at the expense of the thermal energy of the walls. The important point, though, is that it occurs on the approach to thermal equilibrium between the walls and the radiation density in the cavity, not at the stage where equilibrium has already been achieved. However, once thermal equilibrium has been established then the radiation in the cavity will be black.

If an object in a perfectly reflecting cavity absorbs and emits some radiation at all frequencies it is clear that Stewart, Planck and Kirchhoff all held that the full black body spectrum will by achieved by the mechanism outlined numerically above and described by Stewart in the passage quoted. In contrast, Robitaille maintains throughout his series of papers [1–9] that it is necessary to include a black body in the cavity, whether by making part of the walls black or by inclusion of a black object, in order to achieve black radiation in accordance with Kirchhoff's Law.

2.5 Planck's particle of carbon

Robitaille claims that this is precisely why Planck insisted on including a carbon particle in his analysis and why Kirchhoff included one in his experiments. Robitaille dismisses Planck's assertion that the particle merely acts as a catalyst and insists that the carbon particle is responsible for producing the black radiation that Kirchhoff s Law requires. For example, as recently as 2014 Robitaille stated "[Planck's] *cavities all contained black radiation as a direct result* [of placing a carbon particle in the cavity] ... *Since he was driving reflection, all cavities contained the same radiation ...* " [9, p. 158].

However, it is important to distinguish between the *nature* of the black radiation emitted and the *quantity* of it. Planck is perfectly clear that the reason for assuming that the carbon particle is merely a catalyst is that it may be made as small as one likes and, most importantly, its thermal energy can be made so small as to not significantly change the total energy in the cavity [12, § 52]. By definition, therefore, the carbon particle cannot increase the radiation density in the cavity to the level commensurate with the black body temperature; in Robitaille's terms, the particle cannot "drive the reflection", and therefore this cannot be the reason why Planck included it.

Furthermore, if the radiation density is being increased at all frequencies by Stewart's mechanism then there is no need for the particle at all; all one needs to do is wait until thermal equilibrium has been achieved. If the object is a very poor absorber and emitter then this could take some time. In adding a carbon particle to his experiments, Kirchhoff may simply have wanted to accelerate the process.

The situation is somewhat different in the case when the object is a perfect reflector at one or more frequencies. In that case, as Planck stated, the spectrum is black for all frequencies at which the object absorbs and emits but it is indeterminate at the frequencies for which the object is a perfect reflector: *"Hence in a vacuum bounded by totally reflecting walls any state of radiation may persist. But as soon as an arbitrarily small quantity of matter is introduced into the vacuum, a stationary state of radiation is gradually established. In this the radiation of every color which is appreciably absorbed by the substance has the intensity K_ν corresponding to the temperature of the substance and determined by the universal function ..., the intensity of radiation of the other colors remaining indeterminate"* [12, § 51].

However, if the spectrum is indeterminate at any frequencies then it is not possible to properly determine a temperature which is defined in terms of the black body spectrum. See for example *"... the radiation in the new volume V' will not any longer have the character of black radiation, and hence no definite temperature ..."* [12, § 70]. It is apparently in order to avoid this situation that Planck included a particle of carbon which guaranteed that the intensity of radiation was determinate at all frequencies. Why Planck considered that this precaution was necessary is apparent from earlier sections of his work.

Planck had previously discussed the relationship between surface roughness and reflection, pointing out that whether a surface reflected or not was a function of roughness in relation to the wavelength: *"All the distinctions and definitions mentioned in the two preceding paragraphs refer to rays of one definite color only. It might very well happen that, e.g., a surface which is rough for a certain kind of rays must be regarded as smooth for a different kind of rays. It is readily seen that, in general, a surface shows decreasing degrees of roughness for increasing wave lengths. Now, since smooth non-reflecting surfaces do not exist (Sec. 10), it follows that all approximately black surfaces which may be realized in practice (lamp black, platinum black) show appreciable reflection for rays of sufficiently long wave lengths"* [12, § 11].

Thus all objects except perfect black bodies will become reflective at long enough wavelengths. It is apparently in order to avoid this situation that Planck insisted on including a particle of carbon which ensured that all frequencies were present in the equilibrium spectrum. The total radiation energy would not be affected because the particle would not have sufficient energy to do so, by definition. Thus the particle merely acted as a catalyst, as Planck insisted, to convert

the spectrum emitted by the object into a black spectrum as necessary for a proper temperature measurement to be made in accordance with the definition.

Interestingly, despite numerous repetitions in Robitaille's papers [1–8] of his claim that Planck's carbon particle was essential in order to increase the radiation density to the required black body level, Robitaille [9] hints at a change of stance, admitting that *eventually, the cavity might become filled with black radiation, provided that emission and reflection are Lambertian"* [9, p. 160] but then he negates the possibility by stating *"However, for most materials, the introduction of photons into the reflected pool will be inefficient, and the temperature of the system will simply increase. That is the primary reason that arbitrary cavities can never contain black radiation"* [9, p. 160]. In 2015, Robitaille & Crothers [10] return to this theme, stating *"Stewart recognized that, if one could "drive the radiation" in a cavity made from arbitrary materials, by permitting the slow buildup of reflected radiation, the interior could eventually contain black radiation. The argument was true in theory, but not demonstrated in practice"* [10, p. 122].

It appears that Robitaille and Crothers now accept Stewart's mechanism for building up the radiation density by *"bandying about"* the reflection, at least in principle. The authors do not give any explanation for this remarkable *volteface* from Robitaille's earlier works [1–8], but it now appears that his previous objections to Planck's particle of carbon are unfounded: the particle cannot, by definition, increase the total radiation density in the cavity, and Robitaille & Crothers apparently now accept that it is not necessary for the validity of Kirchhoff's Law that it does so.

2.6 Experimental evidence against Kirchhoff's law

Robitaille bases many of his arguments against the validity of Kirchhoff's Law on the fact that black body cavities are never constructed of arbitrary materials; on the contrary, Robitaille insists that manufacturers go to great lengths to construct cavities from special materials to ensure that the radiation is black. Equally, Robitaille points out that resonant microwave cavities cannot contain black radiation. Both these counter-examples are held to demonstrate that Kirchhoff's Law must be incorrect.

However, there appear to be alternative explanations available. In the former case, it may well be that users are concerned about the efficiency of the approach to equilibrium and therefore require black materials in order to speed up the process. It is also likely that manufacturers are concerned, as Planck himself apparently was, to ensure that there are no frequencies at which the cavity is a perfect reflector, which would preclude a proper measurement of temperature.

In the case of microwaves, the cavity is being electromagnetically forced to resonate at a particular frequency and so the radiation cannot be black. Such cases of non-thermal emission were specifically excluded by Plank in deriving his proof: *"A necessary consequence of this is that the coefficient of emission ϵ depends, apart from the frequency ν and the nature of the medium, only on the temperature T. The last statement excludes from our consideration a number of radiation phenomena, such as fluorescence, phosphorescence, electrical and chemical luminosity, ..."* [12, § 7].

Thus it is not logical to conclude that Kirchhoff's Law must necessarily fail because of these supposed counter-examples.

2.7 Challenges to Monte Carlo simulations

Robitaille states that Monte Carlo simulations apparently support Kirchhoff's Law but then he objects on the grounds that: *"Monte Carlo simulations introduce black photons into cavities. Hence, they become black. The process is identical to placing a highly emitting carbon particle, or radiometer, at the opening of a cavity. No proof is provided by computational methods that arbitrary cavities contain black radiation. It can be stated that Monte Carlo simulations obtain similar answers by modeling the repeated emission of photons directly from the cavity walls. In this case, computational analysis relies on internal reflection to arrive at a cavity filled with black radiation"* [5, p. 6].

Apparently, Robitaille's objection to the Monte Carlo simulations is that they rely on Stewart's mechanism for building up the radiation by internal reflection. As Robitaille and Crothers [10] now accept that this mechanism is valid in principle, Robitaille's previous objections to Monte Carlo simulations supporting Kirchhoff's Law should also drop away.

2.8 Super-Planckian emission

Robitaille suggests that recent research into metamaterials supports his arguments. For example, he states: *"Recent results demonstrating super-Planckian thermal emission from hyperbolic metamaterials (HMM) in the near field and emission enhancements in the far field are briefly examined. Such findings highlight that cavity radiation is absolutely dependent on the nature of the cavity and its walls. As previously stated, the constants of Planck and Boltzmann can no longer be viewed as universal"* [9, p. 157].

In relation to the near field emissions, Robitaille refers to three examples from the recent literature [14–16]. All three papers refer to experiments involving bodies with separation distances smaller than the thermal wavelength. However, experimental distances below the thermal wavelength were expressly excluded by Planck: *"Throughout the following discussion it will be assumed that the linear dimensions of all parts of space considered, as well as the radii of curvature of all surfaces under consideration, are large compared with the wave lengths of the rays considered"* [12, § 2].

Planck was concerned about the effects of diffraction at

small scales, in effect limiting his analysis to what are now known as far field effects. Near-field effects are not covered by Kirchhoff's Law and so these three papers cited by Robitaille cannot be used as examples of contraventions of the law. In fact, Guo et al point out that Kirchhoff's Law still suffices to calculate the thermal emission in the far-field and that *"the high-k waves which are thermally excited in the HMM are trapped inside and will be evanescent in vacuum (not reach the far field)"* [14, p. 2]. After comparing the behavior of HMM to other near-field phenomena of surface electromagnetic excitations and photonic crystal structures, Guo et al *"emphasize that in all the above cases including hyperbolic metamaterials, the presence of an interface is enough to guarantee that the far-field emissivity is limited to unity"* [14, p. 5], i.e. that it is Planckian.

The evidence for super-Planckian far-field emissions is not convincing either. Robitaille cites two papers by Yu et al [17, 18] and Nefedov & Melnikov [19] but he notes that Yu et al's claim of emissions in excess of the Stefan-Boltzmann Law made in their arXiv preprint were withdrawn in the published version, and that Nefedov & Melnikov's experiment was not in thermal equilibrium as required by Kirchhoff's Law.

Robitaille's conclusion that *"the universality of blackbody radiation has simply been overstated"* [9, p. 161] does not appear to be warranted on the basis of these recent experiments into metamaterials.

2.9 Robtaille's thought experiment

In [7], Robitaille postulates a thought experiment which he claims disproves Kirchhoff's Law: *"Through the use of two cavities in temperature equilibrium with one another, a thought experiment is presented ... which soundly refutes Kirchhoff's law of thermal emission"* [7, p. 38]. In this thought experiment, the outer cavity is perfectly absorbing and emitting; the second cavity, which is contained entirely within the outer cavity, has perfectly reflecting walls and one side which can be closed remotely. Starting with this inner side open, the two cavities are brought to 4 K; the inner side is then closed; the outer cavity is then heated to 300 K. Robitaille continues: *"The inner cavity walls are thus also brought to 300 K. However, unlike the outer cavity which is filled with blackbody radiation at 300 K, the inner cavity remains filled with blackbody radiation at 4 K. Thereby, Kirchhoff's law is proven to be false"* [7, p. 39].

But by making the inner cavity walls perfectly reflecting and closing the last side, Robitaille has created two entirely separate cavities; by definition, the inner cavity walls cannot emit radiation in either direction, whatever their temperature. They therefore act as boundary walls to what has become a "hollow" outer cavity. The outer cavity no longer contains the inner cavity within itself in the thermal sense; Kirchhoff's Law therefore survives this thought experiment.

3 Robitaille and Crothers 2015 paper

Robitaille & Crothers' paper [10] represents a significant departure from the previous works by Robitaille alone [1–9]. Robitaille and Crothers' *volte-face* on the viability of Stewart's mechanism for filling any cavity with black radiation has been discussed above. However, apart from a re-statement of many of Robitaille's previous objections which have also been discussed above, the thrust of the 2015 jointly-authored paper is to concentrate on criticising Planck's proof of Kirchhoff's Law, a matter only touched on briefly in previous works. Section 4 is titled *"Max Planck and Departure from Objective Reality"* and contains the authors' principal objections to Planck's proof. These will now be examined in detail.

3.1 The meaning of Planck's term "surface"

A number of Robitaille and Crothers' objections hinge on their interpretation of Planck's term "surface" which Planck himself had been careful to distinguish from Kirchhoff's earlier definition. Robitaille and Crothers quote from Planck: *"In defining a blackbody Kirchhoff also assumes that the absorption of incident rays takes place in a layer "infinitely thin". We do not include this in our definition"* [10, p. 124] quoting a footnote from [12, § 10]. In the original text, Planck later explains why he is diverging from Kirchhoff on this point: *"Heat rays are destroyed by absorption. According to the principle of the conservation of energy the energy of heat radiation is thereby changed into other forms of energy (heat, chemical energy). Thus only material particles can absorb heat rays, not elements of surfaces, although sometimes for the sake of brevity the expression absorbing surfaces is used"* [12, § 12]. It appears that Planck could not accept Kirchhoff's "infinitely thin" absorbing layer because it could not include any material particles.

In § 12, Planck is simply being consistent with his earlier discussion of emission: *"The creation of a heat ray is generally denoted by the word emission. According to the principle of the conservation of energy, emission always takes place at the expense of other forms of energy (heat, chemical or electric energy, etc.) and hence it follows that only material particles, not geometrical volumes or surfaces, can emit heat rays. It is true that for the sake of brevity we frequently speak of the surface of a body as radiating heat to the surroundings, but this form of expression does not imply that the surface actually emits heat rays. Strictly speaking, the surface of a body never emits rays, but rather it allows part of the rays coming from the interior to pass through. The other part is reflected inward and according as the fraction transmitted is larger or smaller the surface seems to emit more or less intense radiations"* [12, § 2].

In both § 10 and § 12, it is clear that Planck's use of the term "surface" refers to a geometrical surface dividing two media; the material effects of emission and absorption take

place within the adjoining media. Planck's reference to the surface radiating or absorbing heat is clearly stated as being no more than a convenient shorthand. In contrast, Robitaille and Crothers interpret Planck's term "surface" as being one composed of material particles; it appears that this misinterpretation has led them to a number of erroneous conclusions.

For example, Robitaille and Crothers ask in relation to an element $d\sigma$ of the bounding surface: *"First, what exactly was the location of $\delta\sigma$? In reality it must rest in one of the two media"* [10, p. 127]. This is contrary to Planck's own description of the bounding surface σ as a *"surface separating the two media"* [12, § 35]. Thus Robitaille and Crothers' first objection, that Planck is being inconsistent as to the location of the bounding surface, is unfounded. Similarly, Robitaille and Crothers' second objection to Planck's treatment of the bounding surface, namely *"Planck neglected the fact that real materials can possess finite and differing absorptivities"* [10, p. 127] cannot be maintained.

Robitaille and Crothers raise a third objection to the analysis of an element $d\sigma$ of the bounding surface, namely: *"Third, the simplest means of nullifying the proof leading to Planck's Eq. 42, is to use a perfect reflector as the second medium. In that case, a refractive wave could never enter the second medium and Planck's proof fails"* [10, p. 127]. However, if the surface separating the two media is itself a perfect reflector then the reflectivity on the side of the first medium is obviously equal to 1 but so is the reflectivity for any rays coming from the other side. Thus, $\rho = \rho'$ in accordance with Planck's Eq. 40 leading to his Eq. 42 (see also below) and the proof remains valid. In fact, Planck had already considered this theoretical possibility as occurring for an instant: *"Since the equilibrium is nowise disturbed, if we think of the surface separating the two media as being replaced for an instant by an area entirely impermeable to heat radiation, the laws of the last paragraphs must hold for each of the two substances separately"* [12, § 35]. Obviously the instantaneous nature of this theoretical replacement is necessary to preserve the single system being analysed; a more permanent separation would create two separate systems to which the analysis did not apply. Once again it seems that Robitaille and Crothers' objection is unsustainable.

3.2 Absorption and transmission

Following their quote from Planck's footnote departing from Kirchhoff's definition of an infinitely thin surface in which all the absorption occurred (see above), Robitaille and Crothers commented as follows: *"With his words, Planck redefined the meaning of a blackbody. The step, once again, was vital to his derivation of Kirchhoff's Law, as he relied on transmissive arguments to arrive at its proof. Yet, blackbody radiation relates to opaque objects and this is the first indication that the proofs of Kirchhoff's Law must not be centered on arguments which rely upon transmission. Planck ignored that real surface elements must possess absorption, in apparent contrast with Kirchhoff and without any experimental justification"* [10, p. 124].

However, as is obvious from the passages quoted above, Planck did recognize that absorption must be related to material particles. Once again, the apparent problem arises from the fact that Planck's surface is a geometrical one, whilst Robitaille and Crothers are obviously referring to a surface layer in which, they maintain, all absorption must take place because transmission is not permitted through a black body.

However, Planck also allows for the possibility that absorption in an opaque medium may take place at some unspecified depth below the geometrical surface, i.e. not necessarily in the particles immediately adjacent to the surface. Robitaille and Crothers quote from Planck's description in § 10 of the dependence of the absorbing power on the thickness of the black body material which ends *"The more absorbing a body is, the smaller the value of this minimum thickness, while in the case of bodies with vanishingly small absorbing power only a layer of infinite thickness may be regarded as black"*. Robitaille & Crothers object to this sentence stating that *"Now,* [Planck] *explicitly stated that bodies which are poor absorbers can still be blackbodies. Yet, we do not make blackbodies from materials which have low absorptivities, because these objects have elevated reflectivities, not because they are not infinite"* [10, p. 125] quoting [12, § 10].

But these two objections, about absorptivity and reflectivity respectively, seem to be missing the points that Planck is making: firstly, some absorption may take place by particles situated below the surface. Secondly, Planck had previously stated: *"When a smooth surface completely reflects all incident rays, as is approximately the case with many metallic surfaces, it is termed "reflecting". When a rough surface reflects all incident rays completely and uniformly in all directions, it is called "white". A rough surface having the property of completely transmitting the incident radiation is described as "black"* [12, § 10]. Note that Planck defines black materials as those with a rough surface which does not reflect; all rays falling on a black material pass through Planck's geometrical surface and are subsequently absorbed at some depth in the interior of the black body. No rays are reflected from the body even if the material is, in Planck's terms, a poor absorber. This immediately undermines Robitaille and Crothers' second objection.

Robitaille and Crothers also argue that Planck incorrectly includes transmission within the material of the black body when in fact, Robitaille and Crothers claim, absorption must all occur at the surface: *"Blackbodies are opaque objects without transmission, by definition"* [10, p. 125]. Once again, they are apparently overlooking Planck's definition of a geometrical surface and his careful consideration of where any absorption of radiation passing through that geometrical surface subsequently takes place.

3.3 Reflection

Robitaille and Crothers' § 4.2 deals further with Planck's treatment of reflection. The authors state: *"In the first section of his text, leading to his Eq. 27, ... Planck chose to formally neglect reflection, even though the total energy of the system included those rays which are both emitted/absorbed and those which would have been maintained by driving reflection. Such an approach was suboptimal"* [10, p. 125].

However in the first section of his text, Planck is expressly dealing with the situation within a medium, not with surface effects. His § 25 begins: *"We shall now, as in the previous chapter, assume that we are dealing with homogeneous isotropic media whose condition depends only on the temperature, and we shall inquire what laws the radiation phenomena in them must obey in order to be consistent with the deduction from the second principle mentioned in the preceding section ... Let us consider, first, points of the medium that are far away from the surface"* [12, § 25]. A mathematical treatment then follows, leading to Planck's Eq. 27 towards the end of § 26 which Planck follows with the words *"i.e.: in the interior of a medium in a state of thermodynamic equilibrium the specific intensity of radiation of a certain frequency is equal to the coefficient of emission divided by the coefficient of absorption of the medium for this frequency"* [12, § 26].

Note that Planck is still talking about the interior of the medium where reflection is not applicable because there is no surface; therefore Robitaille and Crothers' objection cannot be maintained.

3.4 Polarization and equality of reflection

Robitaille and Crothers then object to Planck's analysis based initially on a plane-polarised ray, stating: *"In § 5 Planck admitted that homogeneous isotropic media emit only natural or normal, i.e. unpolarized, radiation: "Since the medium was assumed to be isotropic the emitted rays are unpolarized". This statement alone, was sufficient to counter all of the arguments which Planck later utilized to arrive at Kirchhoff's Law [Eq. 42]. That is because the important sections of Planck's derivation, namely § 35–37 make use of plane-polarized light. These steps were detached from experimental reality, relative to heat radiation [Planck, § 35] ..."* [10, p. 127] quoting [12, § 35].

Yet Robitaille and Crothers themselves admit that there was method in Planck's approach, quoting Planck again: *"to prepare for his use of polarized light in later sections, Planck resolved, in § 17, the radiation into its two polarized components"* [10, p. 127], which in itself is unobjectionable. However, Robitaille and Crothers later state that *"such rays could never exist in the context of heat radiation"* [10, p. 129] and this appears to be their principal objection to this means of analysis from which Planck derives the equality of the reflectivity on either side of a geometrical surface separating two different media in his Eq. 40.

But Planck made it clear that an analysis of the special case of polarised light under consideration leads to a valid general conclusion because, as he explained at the end of § 36, the intensity of radiation K_ν, the velocity of propagation q, and the coefficient of reflection ρ at a surface dividing two different media are related by the equation

$$\frac{K_\nu}{K'_\nu}\,\frac{q^2}{q'^2} = \frac{1-\rho'}{1-\rho}\,,$$

where the accented quantities refer to the second medium. Planck continued in §37: *"In the last equation the quantity on the left side is independent of the angle of incidence and of the particular kind of polarization; hence the same must be true for the right side. Hence, whenever the value of this quantity is known for a single angle of incidence and any definite kind of polarization, this value will remain valid for all angles of incidence and all kinds of polarization. Now in the special case when the rays are polarized at right angles to the plane of incidence and strike the bounding surface at the angle of polarization, $\rho = 0$, and $\rho' = 0$. The expression on the right side of the last equation then becomes 1; hence it must always be 1 and we have the general relations:*

$$\rho = \rho' \qquad (40)$$

and

$$q^2 K_\nu = q'^2 K'_\nu \quad (41)".$$

Regarding Planck's Eq. 40, Robitaille and Crothers state bluntly that *"The result was stunning. Max Planck had determined that the reflectivities of all arbitrary media were equal"* [10, p. 129]. On the contrary, what Planck had in fact demonstrated is that the reflectivities on each side of a geometrical surface bounding two different media are equal. Clearly if a different pair of media are chosen, the value of the reflectivity of the bounding surface may be different as well. Planck had previously addressed this point in § 10: *"Since, in general, the properties of a surface depend on both of the bodies which are in contact, this condition shows that the property of blackness as applied to a body depends not only on the nature of the body but also on that of the contiguous medium. A body which is black relatively to air need not be so relatively to glass, and vice versa"* [12, § 10]. Robitaille & Crothers' interpretation that Planck had determined that the reflectivities of all media were equal is unwarranted.

4 Summary and conclusions

Stewart [13] had shown that the radiation in a cavity made from perfectly absorbing material at thermal equilibrium must be black, of an intensity appropriate to the equilibrium temperature. According to Robitaille, Kirchhoff [11] extended this finding to cavities made of arbitrary materials. In a series of papers [1–10], Robitaille has raised various objections to Kirchhoff's extension of Stewart's finding to arbitrary cavities, and to Planck's proof of Kirchhoff's Law [12].

Robitaille concludes that the Law can only be applied validly to cavities containing a black body.

The present paper has investigated Robitaille's claims in depth and compared them to the original papers by Stewart [13], Kirchhoff [11] and Planck [12]. In no instances have Robitaille's objections been found to be sustainable. Furthermore, is has been noted that one of Robitaille's key and often-repeated objections to the build-up of black radiation in an arbitrary cavity according to a mechanism first proposed by Stewart [13] has now been effectively withdrawn in the recent paper by Robitaille and Crothers [10].

Robitaille is obviously correct to point out that black body cavities are never made from reflective materials. However, this fact appears to be more a question of practicality and the need to ensure that the walls are not perfectly reflective at any wavelength so that proper measurements of temperature can be made. It does not seem to amount to a demonstration that Kirchhoff's Law necessarily fails, as Robitaille claims.

This investigation suggests that Kirchhoff's Law and Planck's proof of it remain valid in the situations for which they were intended to apply, including in cavities with walls of any arbitrary materials in thermal equilibrium, unless some other more sustainable objections can be raised in the future.

Acknowledgements

I thank Pierre-Marie Robitaille for encouraging my own research into the subject of Kirchhoff's Law and for engaging in an informative correspondence with me in 2014 regarding Stewart's mechanism for building up the radiation density. I also thank the Mainwaring Archive Foundation for their generous support of my research.

Submitted on: January 8, 2016 / Accepted on: January 11, 2016
First published online on: January 28, 2016

References

1. Robitaille P.-M. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v. 31, no. 6, 1263–1267.
2. Robitaille P.-M. An analysis of universality in blackbody radiation. *Prog. Phys.*, 2006, v. 2, 22–23.
3. Robitaille P.-M. A critical analysis of universality and Kirchhoff's law: A return to Stewart's law of thermal emission. *Prog. Phys.*, 2008, v. 3, 30–35.
4. Robitaille P.-M. Blackbody radiation and the carbon particle. *Prog. Phys.*, 2008, v. 3, 36–55.
5. Robitaille P.-M. Kirchhoff's law of thermal emission: 150 Years. *Prog. Phys.*, 2009, v. 4, 3–13.
6. Robitaille P.-M. Blackbody radiation and the loss of universality: Implications for Planck's formulation and Boltzmann's constant. *Prog. Phys.*, 2009, v. 4, 14–16.
7. Robitaille P.-M. Further insight relative to cavity radiation: a though experiment refuting Kirchhoff's law. *Prog. Phys.*, 2014, v. 10, no. 1, 38–40.
8. Robitaille P.-M. On the equation which governs cavity radiation I. *Prog. Phys.*, 2014, v. 10, no. 2, 126–127.
9. Robitaille P.-M. On the equation which governs cavity radiation II. *Prog. Phys.*, 2014, v. 10, no. 3, 157–162.
10. Robitaille P.-M. and Crothers S.J. "The theory of Heat Radiation" revisited: a commentary on the validity of Kirchhoff's law of thermal emission and Max Planck's claim of universality. *Prog. Phys.*, 2015, v. 11, no. 2, 120–132.
11. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsvermögen und dem Absorptionsvermogen. der Körper fur Wärme und Licht. *Poggendorfs Annalen der Physik und Chemie*, 1860, v. 109, 275–301. (English translation by F. Guthrie: Kirchhoff G. On the relation between the radiating and the absorbing powers of different bodies for light and heat. *Phil. Mag.*, 1860, ser. 4, v. 20, 1–21; DOI 10.1080/14786446008642901).
12. Planck M. Waermestrahlung (1913), Trans. M. Masius as: "The Theory of Heat Radiation". P. Blakiston' Son & Co., Philadelphia, PA, 1914.
13. Stewart B. An account of some experiments on radiant heat, involving an extension of Prévost's theory of exchanges. *Trans. Royal Soc. Edinburgh*, 1858, v. 22, no. 1, 1–20 (also found in Harper's Scientific Memoirs, edited by J. S. Ames: The Laws of Radiation and Absorption: Memoirs of Prévost, Stewart, Kirchhoff, and Kirchhoff and Bunsen, translated and edited by D.B. Brace, American Book Company, New York, 1901, 21–50).
14. Guo Y., Cortez C.L., Molesky S., and Jacob Z. Broadband super-Planckian thermal emission from hyperbolic metamaterials. *Appl. Phys. Let.*, 2012, v. 101, 131106.
15. Biehs S.A., Tschikin M., Messina R. and Ben-Abdallah P., Super-Planckian near-field thermal emission with phonon-polaritonic hyperbolic metamaterials. *Appl. Phys. Let.*, 2013, v. 102, 131106
16. Petersen S.J., Basu S. and Francoeur M. Near-field thermal emission from metamaterials. *Photonics and Nanostructures — Fund. Appl.*, 2013, v. 11, 167–181.
17. Yu Z., Sergeant N.P., Skauli T., Zhang G., Wang H., and Fan S. Enhancing far-field thermal emission with thermal extraction. *Nature Comm.*, 2013, DOI: 10.1038/ncomms2765.
18. Yu Z., Sergeant N., Skauli T., Zhang G., Wang H. and Fan S. Thermal extraction: Enhancing thermal emission of finite size macroscopic blackbody to far-field vacuum. 4 Nov 2012, arXiv:1211.0653v1 [physics.optics].
19. Nefedov I.S. and Melnikov L.A. Super-Planckian far-zone thermal emission from asymmetric hyperbolic metamaterials. 14 Feb 2014, axXiv:1402.3507v1 [physics.optics].

LETTERS TO PROGRESS IN PHYSICS

A Re-examination of Kirchhoff's Law of Thermal Radiation in Relation to Recent Criticisms: Reply

Pierre-Marie Robitaille

Department of Radiology, The Ohio State University, 395 W. 12th Ave, Columbus, Ohio 43210, USA.
robitaille.1@osu.edu

Recently, Robert J. Johnson submitted an analysis of my work, relative to Kirchhoff's Law of Thermal Emission (R.J. Johnson, A Re-examination of Kirchhoff's Law of Thermal Radiation in Relation to Recent Criticisms. *Prog. Phys.*, 2016, v. 12, no. 3, 175–183) in which he reached the conclusion that *"Robitaille's claims are not sustainable and that Kirchhoff's Law and Planck's proof remain valid in the situations for which they were intended to apply, including in cavities with walls of any arbitrary materials in thermal equilibrium"*. However, even a cursory review of Johnson's letter reveals that his conclusions are unjustified. No section constitutes a proper challenge to my writings. Nonetheless, his letter is important, as it serves to underscore the impossibility of defending Kirchhoff's work. At the onset, Kirchhoff formulated his law, based solely on thought experiments and, without any experimental evidence (G. Kirchhoff, Über das Verhältnis zwischen dem Emissionsvermögen und dem Absorptionsvermogen. der Körper fur Wärme und Licht. *Pogg. Ann. Phys. Chem.*, 1860, v. 109, 275–301). Thought experiments, not laboratory confirmation, remain the basis on which Kirchhoff's law is defended, despite the passage of 150 years. For his part, Max Planck tried to derive Kirchhoff's Law by redefining the nature of a black body and relying on the use of polarized radiation, even though he realized that heat radiation is never polarized (Planck M. *The Theory of Heat radiation*. P. Blakiston's Son & Co., Philadelphia, PA, 1914). In advancing his proof of Kirchhoff's Law, Max Planck concluded that the reflectivities of any two arbitrary materials must be equal, though he argued otherwise (see P.-M. Robitaille and S. J. Crothers, *"The Theory of Heat Radiation"* Revisited: A Commentary on the Validity of Kirchhoff's Law of Thermal Emission and Max Planck's Claim of Universality. *Prog. Phys., 2015, v. 11, no. 2, 120–132*). Planck's Eq. 40 (ρ=ρ') , as presented in his textbook, constituted a violation of known optics. Planck reached this conclusion, because he did not properly treat absorption and invoked polarized light in his derivation. Planck also made use of a carbon particle, which he characterized as a simple catalyst. This conjecture can be shown to result in a violation of the First Law of Thermodynamics, if indeed, all cavities must contain black radiation. In the end, while Johnson attempts to defend Planck's proof, his arguments fall short. Though the author has argued that Kirchhoff's law lacks both proper theoretical and experimental proof, Johnson avoids advancing any experimental evidence from the literature for his position. It remains the case that experimental data does not support Kirchhoff's claims and no valid theoretical proof exists.

If a space be entirely surrounded by bodies of the same temperature, so that no rays can penetrate through them, every pencil in the interior of the space must be so constituted, in regard to its quality and intensity, as if it had proceeded from a perfectly black body of the same temperature, and must therefore be independent of the form and nature of the bodies, being determined by temperature alone... In the interior therefore of an opake red-hot body of any temperature, the illumination is always the same, whatever be the constitution of the body in other respects.

Gustav Robert Kirchhoff, 1860 [1]

1 Introduction

Nearly two centuries have elapsed since Gustav Kirchhoff formulated his Law of Thermal Emission [1, 2]. In that time, this law has achieved unquestioned acceptance by the physics community, standing at the very foundation of thermodynamics, condensed matter physics, and astronomy. It constitutes the central pillar upon which Max Planck built his blackbody expression and his claims for universal constants [3, 4]. Eddington's theory of the stars, based on ideal gases, depends on Kirchhoff's law, in order to account for stellar spectra [5]. This remains true for stellar physics to this day [6, 7]. Kirchhoff's law constitutes a citadel for modern astronomy, defend-

ing not only the ideas that stars are gaseous plasmas devoid of lattice structure [5–7], that white dwarfs and neutron stars are highly compressed objects, and that black holes exist [8], but also the concept that a primordial atom once emitted a thermal spectrum and gave rise to the universe [9, 10]. It is precisely because Planck, Eddington, Chandrashekhar, Penzias, Wilson, Dicke, Peebles, Roll, and Wilkinson [1–10] relied on Kirchhoff's law, that they could ignore the central role of the structural lattice in helping to define the emissivity of an object.* While this could be understood in the days of Gustav Kirchhoff, it can no longer be permitted, in light of the tremendous advances made in condensed matter physics and medicine.

Hence, over the course of the past 15 years, I have turned my attention to Kirchhoff's law [13–18, 20–24, 24–26]. My interest in this law did not arise from any desire to study astronomy, but rather, as a consequence of assembling the first ultra high field magnetic resonance imaging (UHFMRI) scanner, at The Ohio State University [27–29]. It was as a direct result of questioning what it meant to say that nuclear magnetic resonance (NMR) and magnetic resonance imaging (MRI) were thermal processes. This had been highlighted long ago by Felix Bloch (Nobel Prize, physics, 1952) who was concerned with thermal processes linking the lattice and the spins [30].†

The laws of emission [1–4, 31–33], are just beginning to impact upon human medicine, as MRI scanners continue to be pushed to ever higher frequencies [27–29]. Thus, there is much more at stake here than the quest to a better understanding of the universe. Correcting Kirchhoff involves moving to a proper description of all thermal processes, not only in physics and astronomy, but in a field as seemingly remote and unrelated as radiology. I have stated that Planck's blackbody law, although valid, remains unlinked to physical reality [12, 17, 19, 23]. That is precisely because of Kirchhoff's faulty law. The physics community has not provided for thermal radiation what is evident for every other spectroscopic process, namely: 1) the setting under which emission occurs, 2) the nature of the energy levels involved, and 3) the nature of the transition species. Only 4) an equation, and 5) the emission of light, have been described [12]. Yet, in every other spectroscopic process, equations are related to physical reality. It takes a hydrogen atom, for instance, to obtain a Lyman or Balmer line. In that case, the transition species is the electron and the electronic orbitals constitute the energy levels. But, for blackbody radiation, spectra are related only to theory, unrestrained by a particular setting, such as the need to have a structural lattice.

That is how astronomers can justify the creation of blackbody spectra from any object. For instance, they have summed a large number of spectroscopic processes to account for the thermal emission from the Sun (see [34] for a complete discussion of this problem). Yet, not one of these processes can be related to the thermal emission from graphite. They have hypothesized that the Big Bang has generated the microwave monopole which surrounds the Earth [10], but have ignored the hydrogen bond from the water which makes up the oceans bathing our planet [35]. Once again, unrestricted by the need to describe thermal emission using a physical mechanism, astronomy has been left to postulate without any consideration of the central physical question in thermal emission: what causes a thermal photon to be emitted by graphite [19]?

Given all that is involved relative to the validity of Kirchhoff's Law [1, 2], Robert Johnson is to be commended, as the first duty of a scientist is to defend established science against possibly false charges. He has also been forthright in submitting a letter to this journal [36], rather than rely on anonymous attacks through social media.

At the same time, it would be an injustice to fail in one's own defense, when a proper understanding of science rests on the outcome. Therefore, I have decided to provide a point by point discussion of Johnson's letter [36]. I do so with the hope that some members of the physics community will begin to take an interest in Kirchhoff's claims and call into question many of the ideas which have been hypothesized [5–10], as a result of concepts which predate the discovery of the atom.

Before I begin analyzing the contents of Johnson's letter, it is vital to outline the setting under which Max Planck viewed a blackbody, as described in *The Theory of Heat Radiation* [4].

Throughout much of his text, Planck make use of perfectly reflecting walls to construct blackbody cavities. As I mentioned previously, this was *"an interesting approach"* [17, p. 4], precisely because such walls, in Planck's context [4], were *"adiabatic, by definition"* [17, p. 4]. They could not participate in generating, or absorbing, a single photon. Moreover, being adiabatic, they were also immune to all conductive and convective processes.

Conversely, unlike Planck, in thinking about perfectly reflecting cavities, I have invoked silver as a nearly ideal reflector of radiation in the infrared [21, 26]. Furthermore, I have insisted that cavities, constructed from such a perfect reflec-

*Nowhere is this more evident than when Eddington insisted that white dwarfs had to possess a small radius, in order to account for their lack of luminosity [5], given the well-established mass-luminosity relationship. Had Eddington considered the critical role of structure in defining emissivity, he would have seen that white dwarfs simply had a different hydrogen based lattice than the hexagonal planar arrangement shared by the Sun and the stars of the main-sequence (see [11, 12] and references therein). But deprived of the use of a lattice, when he stated that all stars could be viewed as ideal gases, Eddington had no other means of explaining the lower than expected luminosity of the white dwarf. Therefore, he was forced to reduce their radius to unreasonable values [5]. This was the first step towards hypothesizing highly dense objects, including the densities now attributed to neutron stars and black holes [8].

†Suffice it to say that the cavity experiments discussed later in this letter have relevance to both blackbody radiation and MRI. Furthermore, any valid analysis of noise power in MRI will be critically based on properly defining and modeling the processes responsible for thermal emission.

tor, possess a characteristic temperature. They are also subject to conductive and convective heat transfer in the establishment of thermal equilibrium. These are important modifications in properly addressing all thermal processes, including radiation, convection, and conduction. For while Planck properly insists that, at thermal equilibrium, there can be *"no conduction"* [4, § 25], no-one maintains that cavities cannot be subject to conductive processes in reaching thermal equilibrium. Laboratory blackbodies are usually brought to temperature using conduction. This will be important later in this letter.

As for adiabatic walls, they could never be characterized by *any* temperature, as I recently emphasized [23]. Consequently, they could never be in thermal equilibrium with anything. Planck stated *"Hence in a vacuum bounded by perfectly reflecting walls any state of radiation may persist"* [4, § 51]. That was very true. But it is also true that such cavities are devoid of any radiation, unless it had previously been injected by some outside means [15–17, 20–23, 26].

In the initial sections of his text, Planck had insisted that all of the energy could be characterized by the radiation field. In truth, the energy must have, at some time, been associated with his oscillators. Otherwise, no photons could have been produced. Thus, Planck's oscillators could be used to produce the field and set thermal equilibrium, but the energy of the system had to be considered as being irreversibly transfered to the radiation field: *"Accordingly we have frequently ... pointed out that the simple propagation of free radiation represents a reversible process. An irreversible element is introduced by the addition of emitting and absorbing substances"* [4, § 170].

This irreversibility and the need for the oscillators to have access to energy, in order to produce the photons, was vital to properly understanding this work. In addition, Planck admitted, in the very last section of his text that *"For the oscillators on which the consideration was based influence only the intensities of rays which correspond to their natural vibration, but they are not capable of changing their frequencies, so long as they exert or suffer no other action than emitting or absorbing radiant energy"* [4, § 190].

Planck insisted that he could place a minute particle of carbon within his cavities. He viewed this object as a catalyst [4, § 51–52], converting radiation within the cavity from one form to another: *"... This change could be brought about by the introduction of a carbon particle, containing a negligible amount of heat as compared with the energy of radiation. This change, of course, refers only to the spectral density of the radiation u_ν, whereas the total density of the energy u remains constant"* [4, § 71]. Planck's particle could only act on the radiation which was already in the cavity. It could not interact with the walls, introduce new energy into the cavity, or set the temperature of the system.

But to interact with the radiation, the carbon particle must have oscillators of its own, functioning over the proper frequency range. Namely, it must be a perfect absorber, characterized by a temperature and part of the thermal equilibrium problem, not a catalyst uncharacterized by any temperature. If devoid of a characteristic temperature, Planck's carbon particle would not contain the proper vibrations to even interact with the radiation in the cavity.

Neither the walls of Planck's perfectly reflecting adiabatic cavities, nor the catalytic carbon particle, could establish temperature. Planck resorted to placing all of the heat within the radiation field. None of the energy could be contained in the walls. He then altered the nature of his walls and removed the requirement that they could not interact with radiation: *"Since, according to this law, we are free to choose any system whatever, we now select from all possible emitting and absorbing systems the simplest conceivable one, namely, one consisting of a large number N of similar stationary oscillators..."* [4, § 135]. Note from this quotation, that Planck could advance no mechanism by which oscillators can actually alter the radiation distribution within the cavity. Planck's oscillators cannot convert the radiation from one form to another, as would be required in the action of Planck's carbon particle were simply catalytic. It remains the case that the radiation contained within a cavity can only be characterized by the nature of the oscillators which produced it. For all these reasons, Planck's carbon particle could never by considered as a catalyst. Indeed, if this particle is attributed with only a catalytic function, it can easily introduce a violation of the First Law of Thermodynamics, as will be seen in § 9 below.

At this point, it is time to address Johnson's submission [36]. In order to maintain the same section numbers, I begin immediately with a review of his introduction [36, § 1].

Johnson's first errors occur in his opening statement, wherein he asserts that I have *"... challenged the validity of Kirchhoff's Law of thermal emission and Planck's derivation of the mathematical form of the universal function of spectral radiance absorbed and emitted by a blackbody"*. There are actually two problems with this statement.

First, I never questioned the mathematical validity of Planck's expression, in the context of an actual blackbody. Rather, I have stated repeatedly that Planck's solution for a blackbody was correct (see e.g. [12, 16, 17, 25]). For instance, in [16, § 1] it is explicitly written that *"The accuracy of Planck's equation has been established beyond question"*. Along with Crothers, I state that *"Fortunately, in Planck's case, the validity of his equation is preserved, but only within the strict confines of the laboratory blackbody"* [25, § 4].*

Secondly, the absorbance of a blackbody does not have a functional form, contrary to Johnson assertion. When Kirch-

*It is troubling that Johnson has misrepresented my position on this matter. My concern has been exclusively centered on Kirchhoff's formulation of a law extending to objects which are not solids and which are constructed from materials lacking a good absorber [13–18, 20–24, 24–26]. I have never questioned the validity of Planck's equation in the case of proper blackbodies.

hoff formulated his law, he defined $E/A = e$ and immediately set A to 1 [1]. This enables the function "e" to have units. Johnson failed to understand, at the onset, Kirchhoff's formulation. Such errors continue throughout his letter [36].

2 Robitaille's earlier papers

2.1 Kirchhoff's law and Planck's proof

Johnson affirms in § 2, relative to Planck's law, that ... *"Planck's derivation is seen as proof of Kirchhoff's law"*. This is not correct.

There are actually several questions addressed in Planck's treatment. This was made clear in the manner in which Planck wrote his book, *The Theory of Heat Radiation* [4].

Planck was primarily concerned with two issues. First, did all arbitrary cavities contain black radiation? This is addressed in the first two chapters [4, § 1–52]. Secondly, Planck was focused on providing the functional form for the blackbody spectrum, through much of the remainder of his presentation. He did so by reviewing the laws of emission advanced by Wein [31] and Stefan [32].* He then discussed Boltzmann and entropy, and presented his oscillators and the blackbody function. In fact, the derivation of the blackbody function itself was completely independent of the derivation of Kirchhoff's law, since when setting $A = 1$, one obtains $E = e$ from Kirchhoff. The functional form of the blackbody spectrum can be obtained, without insisting that all cavities contain black radiation. Planck derived Kirchhoff's Law in the first section of his text solely because of his desire to confer, upon the blackbody expression, universal implications.

If Kirchhoff's Law would be found invalid, as it will eventually become, then Planck does not lose the functional form he supplied describing the radiation of a blackbody, as I have stated repeatedly [12,16,17,25]. However, it would imply that arbitrary cavities are not necessarily blackbodies and that the universality of the constants h and k does not hold [13–18, 20–24, 24–26].

In the next sentence, Johnson writes [36, § 2]: *"However, Robitaille points out that the above definition of Kirchhoff's Law is not complete and furthermore Robtaille maintains that the statement above should be called Stewart's Law as it was originally propounded by Stewart in 1858"*. How could Johnson make such claims?

He begins by omitting an important concept when citing my work. The complete citation is as follows: *"All too frequently, the simple equivalence between apparent spectral absorbance and emission is viewed as a full statement of Kirchhoff's law, adding further confusion to the problem. Kirchhoff's law must always be regarded as extending much beyond this equivalence. It states that the radiation within all true cavities made from arbitrary walls is black. The law of equivalence is Stewart's"* [17]. Importantly, in this citation, I had also included references wherein Kirchhoff's Law was described, solely in the context of the Law of Equivalence, and not within its full scope relative to claiming that a universal function existed. In any event, I never claimed that Kirchhoff's Law was not complete. What I did state was that people often give credit to Kirchhoff for the Law of Equivalence which properly belongs to Stewart [33]. As for Kirchhoff's Law, it is incorrect. Johnson does not seem to understand the fundamental differences between Stewart's Law [33] and Kirchhoff's [1].

The Law of Equivalence [33] simply affirms that, at thermal equilibrium, the radiation emitted by a surface will be equal to the radiation it absorbs, emissivity, ϵ, is equal to absorptivity, α. Stewart did not insist that the radiation inside all cavities was black. That is the reason Kirchhoff's Law [1] does not belong to Stewart [33]. This is an important point, as Johnson falsely asserts, throughout his letter, that Stewart recognized that cavity radiation must always be black. Rather, Stewart recognized that all cavities could become black if they could be driven (see [16] for further discussion). The problem, of course, is that cavities constructed from low emissivity materials cannot be properly driven [15–17, 22–25].

Stewart, while aware of mathematical arguments which might lead to such a conclusion, left the discussion to a footnote [33]. The reason was clear. Stewart recognized, as an experimentalist, that he was not able to prove, in the laboratory, that all arbitrary cavities were black. The experiments described in his work dealt with emission from plates and surfaces [33], not cavities [36]. That was precisely why he did not make a law for cavities, as *The Laws of Physics must be experimentally verified*. In his rebuke of Kirchhoff, Stewart had made the point plainly *"nor did I omit to obtain the best possible experimental verification of my views, or to present this to men of science as the chief feature, grounding theory upon experiment, rather than deducing the experiments from the theory"* (cited in [16]). Stewart never presented any experiments on cavities and therefore, he never made a law related to cavities, as Johnson claims I stated [36, § 2.1].

This was a central difference between the work of Stewart [32] and Kirchhoff [1, 2]. Johnson could have easily come to learn the distinction had he studied the historical review by Seigel [37], which I had cited in [16]. Siegel highlighted that ... *"Kirchhoff himself never performed any experiments which could be construed as attempts at quantitative experimental verification of his law"* [37, p. 588]. Seigel went on to state what Kirchhoff believed: *"... Kirchhoff was rightly pointing out that in this instance neither Stewart's experiments nor his own experiments sufficed to establish a quantitative law, and the burden of the priority claims would therefore have to rest on theoretical proof"* [37, p. 588]. Unfortunately, Kirchhoff was not right. Stewart's experiments were more than adequate to establish the Law of Equivalence. It was with the treatment of cavities that experimental confirmation was lacking.

*Planck never addressed the contributions of Balfour Stewart [33].

In any event, experiments take precedence over theory when it comes to formulating a new law, as our theories are not able to define nature. Furthermore, it is all too easy to accidentally omit a critical element from a theoretical discussion, as has happened when Kirchhoff and Planck unknowingly ignored the energy trapped within the walls of cavities. Such energy can remain forever unavailable to thermal emission. That is why Kirchhoff's Law is invalid. It also provides an illustration of the danger of inferring the laws of physics from theory.

In the end, Seigel also highlighted the difference between Stewart's law and Kirchhoff's claims: *"Stewart's conclusion was correspondingly restricted and did not embrace the sort of connection between the emissive and absorptive powers of different materials, through a universal function of wavelength and temperature, which Kirchhoff established"* [37, p. 84]. It is clear that Stewart's Law did not encompass the universal nature of cavity radiation which Kirchhoff sought, as Johnson attempts to inappropriately claim throughout his letter.

This section closes with Johnson quoting from § 51 of Planck's text [4] and insisting that by placing an *"arbitrary small quantity of matter"* in a perfectly reflecting cavity that *"Planck had thereby demonstrated that all cavities either containing some arbitrary matter, or equivalently having walls made of some arbitrary matter, must also contain black radiation when at thermal equilibrium"*. Yet, in § 51, Planck was placing a small particle of carbon in the cavity. The carbon particle was not an arbitrary material. It was acting as a perfect absorber. I have discussed the inappropriate introduction of a perfect absorber into cavities in detail [16] and will return to the question, once again, in § 2.2, § 2.5 and § 2.9 of this letter.

2.2 Black radiation in a perfectly reflecting cavity

As Johnson opens the third section of his letter, he objects to my conclusion that Planck's statement, *"Hence in a vacuum bounded by perfectly reflecting walls any state of radiation may persist"* [4, § 51], constituted an implicit admission against the validity of Kirchhoff's law.

In trying to defend Planck, Johnson writes: *"However, Planck's statement should perhaps be more properly viewed as a situation to which Kirchhoff's law does not apply because there is no matter present which could absorb or emit radiation."* However, Kirchhoff's law was meant to be independent of the nature of the walls, by definition. Planck associated the temperature of a cavity solely with the radiation it contained, not with any material particles.*. If Kirchhoff was correct, what difference should it make if matter was present to absorb or emit radiation? Nothing in Kirchhoff's law required this restriction and that was precisely the problem. Kirchhoff's law was devoid of all link to actual materials and nature. It was only concerned with hypothetical cavities.

In considering Kirchhoff's law, we can simply examine mathematical limits, as defined by the opaque perfectly absorbing wall (absortivity, $\alpha_\nu = 1$; reflectivity, $\rho_\nu = 0$) and the opaque perfectly reflecting wall ($\alpha_\nu = 0$; $\rho_\nu = 1$). Yet, the second condition led to an undefined expression for Kirchhoff's law, as Planck himself recognized [4, § 48]. It was not possible to claim that a law applies to all materials, when one of its limits was undefined.

Johnson goes on to cite Planck's §51 stating that the radiation within all cavities will always be black, even though Planck, in the very same section, has just introduced a particle of carbon in this cavity, which Johnson recognizes as being a *"perfect absorber and emitter at all frequencies"* [36]. But, Planck viewed the carbon particle as a catalyst [4, § 51–51]. Johnson then writes, in speaking of Planck: *"Note that the quoted statement covers both the situation where the object absorbs or emits over all frequencies, and the situation where some frequencies are not absorbed or emitted at all"* [36].

Planck reached his conclusion by inserting a particle of carbon. This ensured absorption and emission over all frequencies. Planck never demonstrated that this applied to *situations where some frequencies are not absorbed or emitted at all"*, as Johnson claims [36]. Planck placed the carbon particle within the cavity and then claimed that it acted only as a catalyst. He sidestepped the fact that this particle was acting as a perfect absorber, and thereby controlled the entire problem. I have already demonstrated this fact mathematically and the reality that arbitrary cavities, at thermal equilibrium, do not contain black radiation [15]. Importantly, Johnson's letter fails to address these simple algebreic proofs that Kirchhoff's Law cannot be valid [15].† Again, I will return to the question of the carbon particle in § 2.5 and § 2.9.

2.3 The approach to equilibrium

Johnson opens this section by pondering what was correct: Do all cavities contain black radiation, as Kirchhoff and Planck held, or do arbitrary cavities contain arbitrary radiation, as Robitaille asserted? The question was simple enough to answer, as blackbodies are always constructed from good absorbers.

In fact, had Johnson considered the history of blackbody radiation, he would have recognized that arbitrary cavities are never black. That is why those who provided Max Planck with the data used to verify his equation worked so hard to

*"Still, even Planck recognized that material objects were required to establish a temperature, *"But the temperature of a radiation cannot be determined unless it be brought into thermodynamic equilibrium with a system of molecules or oscillators, the temperature of which is known from other sources"* [4, § 144]

†Reference [15] contains a detailed analysis of some of the problems with Kirchhoff's logical arguments in advancing his proofs [1,2]. It also contains simple proofs of Stewart's Law of Equivalence [33] and clear demonstrations that arbitrary cavities, under conditions of thermal equilibrium, are not black. Johnson cannot ignore these proofs in his letter, if he wishes to honestly evaluate my work.

construct laboratory blackbodies which provided the proper functional form [38–40]. These papers, especially the review by Hoffmann [38], are important to study, because they highlight the complexity of building proper blackbodies.

As a simple example, the problem can be viewed to involve, to some extent, the behavior of graphite itself. In the visible range, some forms of graphite, which are mined, can be relatively good absorbers, but others, surprisingly, can be rather poor, as can be ascertained by examining emissivity tables [41]. However, as one becomes increasingly interested in the region towards the infrared, graphite begins to fail. This has been known since the days of Langley at the end of the 19th century [16, § 2.1]. That is why materials like the metal blacks are utilized, in this region of the electromagnetic spectrum, to assemble blackbodies [42–45].

We already have the experimental proof, but most people simply ignore these laboratory realities. For, if Kirchhoff's Law was valid, there would be no need for metal blacks in building laboratory blackbodies and German scientists would not have used rolled platinum and specialized mixtures of chromium, nickel, and cobalt oxide to blacken the interior of their cavities [38, p. 57]. Such mixtures indicate that their was nothing arbitrary in the construction of blackbodies.

This remains a specialized field and such objects are *always* sophisticated devices unavailable when Kirchhoff extended his law to all cavities.*

In this same section of his letter, Johnson goes on to consider what would happen to the radiation, within an arbitrary cavity, if the initial radiation was less than the maximal hypothesized by Kirchoff's law. The arguments he advanced are flawed at a fundamental level.

Johnson first places an opaque object within a perfectly reflecting cavity and defines that the intensity of the radiation is 100 within the cavity, the proper value for black radiation. He assumes that the object has an emissivity of 0.8 and then states that when radiation within the cavity interacts with the object, 80 units will be absorbed/re-emitted and 20 units being reflected. Johnson notes that the radiation within such a cavity will remain black at 100 units. Of course, the experiment is false, as an object with an emissivity of only 0.8 could never fill the cavity with black radiation in the first place. The radiation would have to be increased by some other means.† Deviations from this case are only permitted if thermal equilibrium has been violated, after the cavity and the object reached the temperature of interest, or if the perfectly reflecting cavity has otherwise been filled with black radiation [16, 17]. It is important to recall, that even the sampling of a cavity with a detector can act to fill it with black radiation [17, § 2]. Therefore, this situation, as described by Johnson, does not lend any support to Kirchhoff's claims. It was simply ill-conceived.

At this point, Johnson considers another scenario wherein an object with an emissivity of 0.8 can only emit 80 units initially into the cavity. These 80 units then strike the wall and reflect back towards the object, where now he claims that only 64 units are absorbed (since the emissivity is 0.8), and 16 units are reflected. Johnson notes that the object *"... was bound by its initial temperature to continue emitting 80 units"* [36, § 2.3]. He notes the shortfall in the total amount of radiation absorbed by the object, and claims that this can only be rectified by lowering the temperature of the object. The errors in logic are striking.

First, Johnson fails to recognize that it is the total radiation coming off the object at thermal equilibrium which matters. That total radiation is equal to 64 units emitted and 16 units reflected at the onset, because the cavity and the object are already at thermal equilibrium, by definition. Johnson does not get to say that the object must emit 80 units to begin his experiment and then state that only 64 units are absorbed and re-emitted. He can only sample the total radiation coming off the object. He has no means of distinguishing what was, in fact, reflected and what was emitted. He only knows that 80 units came off the object. These are then reflected off the wall and travel towards the object, where 64 units will be absorbed, then re-emitted, while 16 units will be reflected. Johnson also fails to understand that he cannot allow the temperature of the object to drop, as this is a violation of the zeroth law of thermodynamics. For my part, I would not disallow thermal equilibrium between the cavity and the object, as Johnson asserts.‡

Relative to the last experiment, it is interesting to note what Johnson has actually done. At first, he ignored reflection, stating that all 80 units leaving the object were emitted. Then, on absorption, he now considered reflection, permitting only 64 units to be absorbed and the remaining 16 to be reflected. So what has happened?

Note, for instance, that when Max Planck derived the first section of his proof of Kirchhoff's Law, he also ignored reflection (see [25, § 4.2] for a complete description of what Planck did in this instance). Robitaille and Crothers note that Planck was allowed to ignore reflection, as these terms, if retained, could be canceled out [25, § 4.2]. They also demonstrate that the full treatment retaining reflection can lead to additional insight, relative to this problem [25, § 4.2].

If Planck was allowed to ignore reflection, perhaps this can be most easily explained by examining the Law of Equivalence itself [33]. I have already highlighted that Stewart's Law can be written either as, $\epsilon_\nu=\alpha_\nu$, or as, $\epsilon_\nu+\rho_\nu=\alpha_\nu+\rho_\nu$ [15]. The use of either form will lead to the correct answer. However, what Johnson has done was to mix the two forms of Stewart's law, inventing a scenario wherein he sets $\epsilon_\nu=\alpha_\nu+\rho_\nu$,

*The author has reviewed laboratory blackbodies in [16, 17].

†I have already demonstrated mathematically, that the radiation in the cavity, in this case, will not be black but will have an intensity appropriate for the emissivity of the object it contained [15].

‡I also reject all of Johnson's other deductions relative to how I would view an experiment which I never even described in my papers.

which is clearly false.

At this point, Johnson once again tries to state that I have attributed Kirchhoff's Law to Balfour Stewart. In this, he misses the central point. Stewart's footnote does not make a law of physics. It presents a mathematical argument. Stewart recognized that, if he wanted a blackbody spectrum from a cavity, he must have recourse to lampblack. Johnson believes that Stewart was specific on this point, arguing for a *"theoretical leap"* [36]. But in so thinking, he failed to recognize what Stewart understood: cavities can only be *demonstrated* to be black experimentally if they contain a good emitter. Stewart did hypothesize extensively about bandied radiation, well after 1858 (see [16] for a complete discussion), and conjectured that cavities of low emissivity can be made to appear black. The thesis was never proven and with good reason [15–17, 22–25]. Stewart stated a theoretical idea, not a law. The point has been made clearly by Seigel, as noted in § 2.1 above [37]. That is why I wrote in my initial paper: *"Stewart realizes that the lampblack surface within the enclosure is essential"* [16]. Stewart might have had a theoretical argument, but he did not have data. It is in this aspect that he was much more prudent than Kirchhoff when he presented his work [33]. That is why I have always acknowledged this Scottish scientist. Stewart exercised wisdom in 1858 [33] and Johnson shall not deprive him of this quality.

2.4 Stewart's treatment of reflection

Johnson then goes on to describe, in detail, Stewart's footnote, as if this was central to the idea which Stewart was conveying. Stewart's paper deals with the Law of Equivalence, not with cavity radiation and universality [33]. The argument which Johnson resurrects is contained in a footnote, precisely because this constitutes its proper position in the paper. Stewart makes us aware that he understands a mathematical argument previously advanced by others (see references contained in [17]), but he does not raise them to a central part of this thesis, because these ideas were not supported by laboratory data.

In considering the bandied radiation, Johnson makes the claim that the energy required to fill the cavity can be extracted from the walls in order to drive *"Stewart's mechanism"*. In this aspect of his letter, Johnson is actually repeating ideas from my own papers on cavity radiation, wherein such processes have already been discussed in detail [21, 23, 26].* Johnson adds nothing new to this discussion. He also fails to understand, at a fundamental level, that it is by invoking the energy retained in the wall that Kirchhoff's Law can be proven to be false [21, 23, 26]. Planck specifically used an adiabatic wall which could not be characterized by any temperature to build his perfect reflector, because he wanted all of the energy to be contained in the field, not in the wall [4]. Since adiabatic walls are detached from all thermal processes (i.e. radiation, conduction, convection), they cannot be characterized by any temperature [21].

Johnson analyzes Stewart's experiments [33] with low emissivity plates in obtaining the same functional form as if the plate had been black. Yet, it is not solely a question of time elapsed, as he attempts to argue. For instance, he permits the temperature of one of his plates to drop in clear violation of the Zeroth Law of thermodynamics *"... the only difference in this case is that during the initial period the partially absorbing plates is absorbing less radiation than it is emitting; it is therefore cooling down and part of its initial energy is being used to increase the radiation density between the plates, or, in Robitaille's terms, in "driving the reflection"* [36, § 2.4]. I never permitted an object temperature to drop, in order to drive the reflection.

My papers are concerned with a law defined under thermal equilibrium, not the approach to equilibrium. I have highlighted that one cannot create photons from nothing. Scientists are not permitted to violate the First Law. What has happened in this letter is that Johnson permits the temperature to drop in order to avoid violating the First Law, as he knows that he must get photons from somewhere. The arguments are all invalid, as we are concerned with a system in thermal equilibrium, not the approach to such equilibrium.

While Johnson understands that the idea of driving a cavity is an important concept, he continues to ignore its consequences. For instance, such processes rely on access to a perfect absorber, or some temporary violation of thermal equilibrium [15–17, 22–25]. They are also prone to introduce a violation of the First Law of Thermodynamics, as energy must come from somewhere. Also, energy cannot be destroyed.

The question, relative to thought experiments, relates to the origin of the energy entering a cavity once it is already at thermal equilibrium. Provided that the cavity walls are not adiabatic, but can be represented by graphite, or silver, then there are three scenarios to consider: 1) energy enters the system from outside, 2) energy travels reversibly out of the walls of the cavity to irreversibly fill the cavity [21, 26], and 3) energy is irreversbilty trapped within the cavity walls [26]. None of these possibilities have ever been considered by Planck. They arise from the assembly of work which is currently being challenged by Johnson.

Let us assume that the energy came from outside the system. Then, once it reaches the cavity walls, it must be allowed to either 1) help fill the cavity with additional photons, or 2) dissipate additional energy into the walls of the cavity. However, the cavity walls are already at a given temperature. To permit additional energy to enter would alter this value. As such, no energy can be allowed to enter the walls, as this would violate the zeroth law. Thus, if any energy enters the cavity walls from outside the system, it must simultaneously leave and produce additional photons in the interior. It is clear that, with such a scenario, if the walls are fully reversible

*The author published [26] just a few days before Johnson submitted his letter and he was made aware of this work.

stores of energy, the cavity will become filled with radiation. The problem becomes, when does one stop? Obviously, the experimentalist can place any amount of photons in the cavity, given enough available energy from the outside and no concern for the First Law. If the radiation intensity within the cavity becomes too great, then he can simply affirm that thermal equilibrium has been violated and that the cavity must now be represented by a higher temperature.

As for the idea that the energy contained within the walls can be reversibly used to fill the cavity with even more radiation, I have already considered the concept on two occasions [21, 26]. In reality, such processes are likely to be physically impossible. Thermodynamically, the concept is allowed, but the problem is that, if the energy of the walls if fully available to build up photons in the interior, the cavity would already be black, unless specialized means are used to isolate this energy [26]. In reality, every material which is not a perfect emitter will actually possess at least some energy which is irreversibly trapped in the walls [26] relative to the ability to support emission. That is the central reason why arbitrary cavities are never black. Planck had considered that only the production of the radiation field was irreversible, as I discussed in the introduction. This may have been everyone's major stumbling block relative to cavity radiation and Kirchhoff's Law. Prior to 1906, when Planck's lectures where written [4, p. xi], neither he, nor Kirchhoff, understood that some of the energy which enters a metal will be trapped in its conduction band electrons and forever remain unavailable to emission [26].* We shall return to *"Stewart's mechanism"* in § 2.5, § 2.6, § 2.7, and § 2.9.

2.5 Planck's particle of carbon

Johnson then moves to try to defend Planck's use of a carbon particle as a simple catalyst. I have already spoken extensively on this issue: Planck's carbon particle is not a catalyst [16]. It is a perfect absorber/emitter. Planck uses carbon, not a particle of some other material, and with good reason. He needs a perfect absorber. It is not simply a question of having a particle which can absorb over all frequencies of interest. In fact, a quick study of emissivity tables would demonstrate that, if this were the case, Max Planck had many other materials available to him [41]. He wanted a perfect absorber and, when he placed it in his cavity, as I have said previously, it was as if he had coated the entire inner surface with lampblack. Otherwise, what does it mean to be "perfect"? As I stated in the introduction, the reality remains that Planck's carbon particle must have access to oscillators, otherwise it cannot even interact with the radiation. It must also be characterizable with a temperature, such its oscillators could operate over the entire range of frequencies required to make the cavity radiation black at the proper temperature. The need for this temperature directly implies that the carbon particle is a perfect absorber, not a catalyst.

Johnson claims that there is a difference between *"the nature of the black radiation and the quantity of it"*. He then argues that Planck has made the particle small such that its energy content can be neglected relative to filling the cavity with radiation. Planck's position and Johnson's defense are not well-reasoned in that they neglect that the particle and cavity must be allowed to come to thermal equilibrium. This is one of the reasons why Planck's use of an adiabatic wall to build a perfectly reflecting cavity is not appropriate. Planck also attempted to deprive the carbon particle of a specific temperature. In so doing, he was overlooking the very detail which was critical to obtaining the proper answer (see also § 2.9). Johnson states, *"... By definition, therefore, the carbon particle cannot increase the radiation density in the cavity to the level commensurate with the black body temperature; in Robitaille's terms, the particle cannot "drive the reflection", and therefore this cannot be the reason why Planck included it. Furthermore, if the radiation density is being increased at all frequencies by Stewart's mechanism then there is no need to include the carbon particle"* [36, § 2.5]. Unfortunately, for Johnson, he cannot resort to *"Stewart's mechanism"*, as he cannot practically demonstrate its validity in the context of a perfect reflector. Even Stewart, cannot generate photons from a perfectly reflecting cavity. The issue at hand is the carbon particle, not *"Stewart's mechanism"*.

As such, let us first consider the proper way of viewing the carbon particle, then return to Planck and Johnson, both in this section and in § 2.9.

The simplest means of addressing this problem is to consider that whatever light is reflected off the walls of a perfectly reflecting cavity can strike the particle. The particle must then transform the radiation and return this light back towards the cavity walls [15]. The temperature of both the cavity and the particle must be the same and the temperature of the latter must not be allowed to drop in order to respect the zeroth law. Under this condition, full equilibrium between the walls and the particle would exist and the cavity could easily be demonstrated to contain black radiation [15]. Herein was the power of equilibrium arguments.

In order to further clarify the point, let us consider what was physically occurring within the cavity when Planck introduced his small particle of carbon. Since the cavity was perfectly reflecting, we can assume that it can be best approximated by polished silver [23,26], not by an adiabatic wall [4]. The emissivity of the cavity must be 0 and it initially contains no photons. Let us surround the cavity with an adiabatic wall, in order to isolate the system.

As a result, the temperature of the cavity in this case is defined by the energy content of its walls. When the carbon particle is introduced into such a system, even if it contains no appreciable heat on its own, it also comes into thermal contact with the wall of the perfectly reflecting cavity. At this point,

*The energy can still be removed from the wall through conduction and convection.

thermal energy will become available to the carbon particle from the cavity wall. This particle can then transform the energy which would be otherwise irreversibly trapped in the walls [26] and fill the cavity with radiation. In this sense, the carbon particle acted as a transformer, converting phonon energy and/or energy associated with thermal conduction in the silver wall into photons. It was not a catalyst, as it was critical to conversion occurring. I have always modeled perfect reflectors using silver [23, 26], not using adiabatic walls. Without the carbon particle, the cavity would remain devoid of any radiation and all of its energy would remain forever trapped in its walls.

As for the case considered by Max Planck, an adiabatic wall contained no energy. Therefore, the carbon particle, devoid of significant heat, could never fill such a cavity with radiation at any temperature.

Contrary to Johnson's claim, neither Crothers, nor I, have said that the carbon particle cannot increase the radiation inside the cavity. Rather, my papers provided the only means for the carbon particle to fill the perfectly reflecting cavity. As for Johnson, he must adopt Planck's position, and remain forever unable to consider the content of the walls and the ability of the carbon particle to transform this energy into photons. He cannot be permitted to jump between my model and Planck's, as this is the entire basis of this discussion.

If Planck stated that *"Hence in a vacuum bounded by totally reflecting walls any state of radiation may persist"* [4, §51], it was because he recognized that Kirchhoff's law became undefined when A=0. But that does not mean that the cavity in this case contains forms of radiation which are blackbody, unless such radiation has been introduced by some outside mechanism.* In fact, the perfectly reflecting cavity must be considered empty, because it had no means of producing a photon and all of its energy content was trapped in its walls before the introduction of the carbon particle [26]. Johnson argues that, if the spectrum is indeterminate at any frequency, it is impossible to set a temperature. Again this is false, as the walls also contain energy [26]. Max Planck also ignored this fact, a critical error in selecting adiabatic walls.

Johnson then cites Planck's discussion [4, § 11] that all objects show significant reflection at sufficiently long wavelengths, except perfect blackbodies. He concludes that this is why Planck introduced the carbon particle [36]. But, if that was true, then Planck's introduction of the carbon particle would be acting to make all cavities perfect blackbodies, a point which supports my position.

In closing this section, Johnson makes the charge that I now accept, at least in principle, *"Stewart's mechanism"* for building up the reflection within a cavity. He alleges a remarkable *"volte-face"* on my part when I published a paper with Crothers [25].

Such a conclusion is not reasonable, as my papers have always considered Stewart's hypothesis (see e.g. [15–17, 22–24] all of which precede [25]). Furthermore, Crothers and I have restated, in no uncertain terms, that *"Stewart's mechanism"* does not work [25, § 2, 3].

Over the years, I have dealt consistently with the problem of thermal emission and have always held the position that arbitrary cavities are not black. I have examined numerous questions including 1) perfectly absorbing cavities, 2) perfectly reflecting cavities, 3) perfectly reflecting cavities containing a carbon particle, 4) perfectly reflecting cavities containing an arbitrary object, 5) perfectly reflecting cavities containing two arbitrary objects, 6) perfectly absorbing cavities containing a perfectly reflecting cavity, 7) two cavity problems (both for the reversible and the irreversible cases), 8) actual laboratory blackbodies, 9) Kirchhoff's two faulty initial proofs, 10) Planck's faulty proof of Kirchhoff's Law, 11) proper equations governing cavity radiation and 12) effects of driving the reflection term. Nowhere have I ever stated that *"Stewart's mechanism"*, as Johnson refers to bandied reflection, can ever lead to black radiation in all cavities, despite repeatedly addressing the question [15–17, 22–25]. What I have stated is that, if one tries to drive a cavity made from materials with a low emissivity, in order to build up black radiation in its interior, it is likely that the cavity will simply prefer to move to a higher temperature [23]. That is because any energy introduced into the cavity must also be available to the walls. If those walls cannot easily emit a photon, they will simply increase their temperature. Moreover, I have emphasized that the use of bandied radiation, even if possible, could only lead to filling a cavity with black radiation, in the ideal that the walls were capable of Lambertian reflection [23]. No specular reflection must have taken place and all reflection must have been diffuse. Otherwise, one risks generating standing waves, as I have previously highlighted [16] (see also § 2.6). Johnson ignores all these points when he addresses bandied radiation.

2.6 Experimental evidence against Kirchhoff's law

This is perhaps the most unusual section of Johnson's letter [36], as he tries to explain why manufacturers do not build blackbodies from arbitrary materials. Rather than concede that this constitutes direct experimental evidence against Kirchhoff's law, as I have stated, Johnson reaches for the indefensible. He argues: *"It is also likely that manufacturers are concerned, as Planck himself apparently was, to ensure that there are no frequencies at which the cavity is a perfect reflector, which would preclude a proper measurement of temperature"*.

Kirchhoff's Law demands that all cavities be black, independent of the nature of the walls. Manufacturers are not concerned with materials acting as perfect reflectors, since most solids emit continuous spectra over a wide range of frequencies. The problem is that many solids are poor emitters,

*I will return to this issue in § 2.9.

not that they are perfect reflectors.

Furthermore, the temperature of a cavity in the laboratory is determined by temperature sensors in its walls. Cavities are heated, conductively or otherwise, the temperature on the sensors in its walls are noted, and thermal equilibrium is defined by those sensors maintaining a stable temperature reading. Establishing the temperature of a laboratory cavity has nothing to do with measuring its radiation field and it would be irrelevant, if some frequencies were absent from the spectrum. This cannot affect the reading of a sensor in the wall of the cavity.* Even Planck recognized that a proper measure of temperature depends on the use of sensors or thermometers: *"But the temperature of a radiation cannot be determined unless it be brought into thermodynamic equilibrium with a system of molecules or oscillators, the temperature of which is known from other sources"* [4, § 144].

Johnson then moves to question any work in microwave cavities. He launches this new challenge precisely because these cavities are known not to contain black radiation, as I have demonstrated experimentally using UHF frequencies near the microwave region [17]. In attempting to dismiss microwave cavities, Johnson cites Planck: *"The last statement excludes from our consideration a number of radiation phenomena such as fluorescence, phosphorescence, electrical and chemical luminosity"* [4, § 7]. Johnson's use of such a quotation relative to microwave cavities demonstrates that he does not fully understand the experimental problem.

Kirchhoff's law allows for the presence of any object within the cavity. Therefore, the resonant elements used in my own work [17] are allowed, as they do not emit a single photon. They build up standing waves. Still, it remains clear that fluorescence, phosphorescence, electrical and chemical luminosity cannot be considered.†

The microwave cavity is not producing radiation by some non-thermal means, like fluorescence, phosphorescence, electrical or chemical luminosity. Rather, it is being subjected to sampling by a network analyzer which is sending microwave energy into the enclosure and noting what energy returns. If the cavity is able to reflect some of this radiation internally, then it can build up standing waves. Alternatively, if the cavity is truly black, then it should be able to absorb all the energy coming from the network analyzer with no returned energy. In any case, such return-loss measurements on cavities are routinely done throughout thermometry (see references cited in [48]).

In the infrared, cavities can be subjected to radiation from a standard blackbody, for instance, in order to verify their absorptivity by noting the returned energy.‡ In the microwave, when testing blackbodies for satellites, the source is often a network analyzer (see references cited in [48, 49]). This is a common measurement in testing the quality of blackbodies at these frequencies.

Johnson must recognize that microwave cavities are utilized on satellites such as COBE [50] and PLANCK [49]. These cavities are tested using return-loss methods, exactly as I have done in [17], when testing an MRI cavity. Many of these cavities are not black, including some which have been claimed as such and launched aboard satellites [49]. Microwave cavities often contain signs of standing waves, as radiation from the network analyzer enters the cavity. The presence of such standing waves provides solid evidence that not all cavities in the microwave contain blackbody radiation. This is an important point to recognize, as Johnson would like to build up arbitrary radiation in cavities with reflection, using *"Stewart's mechanism"*. Standing waves demonstrate that the presence of specular reflection within a cavity is always counter to the interior containing black radiation [48,49]. This highlights yet another problem with *"Stewart's mechanism"*. It is critically dependent on any reflection within a cavity being diffuse and not specular. Otherwise, the radiation will not be Lambertian, as required of a blackbody.

Contrary to Johnson's position, experiments with cavities in MRI provide strong evidence that Kirchhoff's law does not hold (see [17] and references therein). This is especially true given that Kirchhoff's Law has been generalized to treat geometries where diffraction becomes important (see [17] and references therein). Furthermore, microwave studies demonstrate that small cavities, containing only a few centimeters of Ecosorb and conductively anchored to a radiation shield, like the 4K reference loads on the Planck satellite, can never be black [48, 49]. This presents a serious problem for those interested in the LFI data produced by this satellite [49].

Once again, the fact remains that Kirchhoff's Law does not have any valid experimental support. Arbitrary cavities are not black and this reality has consequences which must not be ignored.

*However, for real blackbodies, when the temperature sensors indicate a certain temperature, one can be assured that the radiation sampled will be black.

†Surprisingly however, in Kirchhoff's initial paper [2] he actually insists that even fluorescent material could be included within the cavity and it will still be black: *"It may be observed, by the way, that the proposition demonstrated in this section does not cease to hold good even if some of the bodies are fluorescent. A fluorescent body may be defined as one whose radiating power depends on the rays incident on it for the time being. The equation E/A = e cannot generally be true for such a body; but it is true if the body enclosed in a black covering of the same temperature as itself, since the same considerations that led to the equation in question on the hypothesis that the body C was not fluorescent, avail in this case even if the body C be supposed to be fluorescent.*" [2]. These arguments are removed however, without explanation, when Kirchhoff's work is revised several years later [46]. Still, this indicates a flaw in Kirchhoff's initial derivation of his law [2], as he had thought that his derivation applied to fluorescent bodies, which was not correct. There are indeed flaws in Kirchhoff's initial derivation, as the author has independently ascertained [15]. Moreover, Schirrmacher [47] has reviewed the proofs of Kirchhoff's Law before and after Planck [48]. Even in 1912, Hilbert complained that a valid proof a Kirchhoff's law still did not exist [47], even though the Planck's lectures on the subject were given in 1906 [4, p. xi]. Such a proof is lacking, to this day.

‡Note that if Kirchoff's law was correct, there would be no need to have standard blackbodies in order to calibrate other cavities, as all cavities would be black.

2.7 Challenges to Monte Carlo simulations

Johnson then moves to briefly discuss Monte Carlo simulations in a single paragraph stating: *"Apparently, Robitaille's objection to the Monte Carlo simulations is that they rely on Stewart's mechanism for building up the radiation by internal reflection. As Robitaille and Crothers now accept that this mechanism is valid in principle, Robtaille's previous objections to Monte Carlo simulations supporting Kirchhoff's Law should also drop away"*.

Clearly, I have never accepted *"Stewart's mechanism"* for building up radiation within a cavity. First, such a mechanism, under certain circumstances, constitutes a violation of the First Law of Thermodynamics. Secondly, it is not possible to place energy into the interior of a cavity without also potentially placing energy into the walls. This is never considered by Monte Carlo simulations, and that is why they remain invalid. Such simulations agree with Kirchhoff Law, precisely because they ignore the dynamics going on in the wall and *a priori* forbid the temperature of the wall to rise in lieu of emitting a new photon.

2.8 Super-Planckian emission

Johnson's letter then examines my treatment of metamaterials [23]. He argues that Planck specifically excluded the near field by quoting: *"Throughout the following discussion it will be assumed that the linear dimensions of all parts of space considered, as well as the radii of curvature of all surfaces under consideration, are large compared with the wave lengths of the rays considered"* [4, § 2]. On the surface, this is a good point. Planck is clearly allowed to restrict his derivation. This does not mean, however, that the near field region cannot be considered today, in order to shed additional light on thermal emission.

In this regard, Kirchhoff's law has been generalized to include the limit initially excluded by both Kirchhoff and Planck [17, § 3]. The near field behavior can be considered for additional insight and the point raised by Johnson is weak at best. Science does get to move forward.

Johnson then goes on to claim that the evidence in the far field, is not convincing. He notes from Guo et al. [51] that: *"the presence of an interface is enough to guarantee that the far-field emissivity is limited to 1"* [36]. Guo's statement is noteworthy. However, Johnson neglects to cite the following from Guo's paper: *"The usual upper limit to the blackbody emission is not fundamental and arises since energy is carried to the far-field only by propagating waves emanating from the heated source. If one allows for energy transport in the near-field using evanescent waves, this limit can be overcome"* [51].

It is clear that the study of metamaterials is an area of science which is just beginning to be explored. It is also not established that far-field behavior will always adhere to the limits set forth by Planck's law. This is why I previously highlighted [23] the work by Yu et al. [52] and [53]. Yu et al. removed the claim made in the arXiv version of their paper [53] when they published their *Nature Communications* paper [52]. Here is the exact quotation from my paper on this issue: *"In that case, the spatial extent of the blackbody is enhanced by adding a transparent material above the site of thermal emission. A four-fold enhancement of the far-field emission could thus be produced."* In their *Nature Communications* article, the authors argue that this does not constitute a violation of the Stefan-Boltzmann law, because the effective "emitting surface" is now governed by the transmitter, which is essentially transparent. However, this was not the position advanced when the results were first announced and the authors wrote: *"The aim of our paper here is to show that a macroscopic blackbody in fact can emit more thermal radiation to far field vacuum than* $P = \sigma T^4$*"* [53].

In Yu's work, the emission is arising from a small blackened disk of material [52, 53]. The photons emitted from this surface greatly exceed anything predicted by Planck. At issue is the assignment of the emitting surface, from a theoretical perspective. Is it the blackened disk, which is the only possible source of photons, or the transparent shield? The key difficulty for blackbody radiation science is that blackbodies were always defined as opaque objects. Hence, it is difficult to conceive why the blackened disk should not be considered as the proper emitting surface in this problem. But assigning the emission to a transparent surface is now the only way of salvaging Kirchhoff's law. Once again, note how Kirchhoff had worded his law in the quotation at the very beginning of this reply. He was referring to opaque objects.

Then, there is the problem that, during Yu et al's experiment, the blackened disk is always heated [52, 53]. This implies that thermal equilibrium does not exist, since conduction of energy, which is heating the disk, must be considered. However, if this is to be used as an argument against these findings, then what of the problem of continuously heating ordinary cavities, in order to maintain their temperature equilibrium? As I previously stated: *"Obviously, modern experiments fall short of the requirements for thermal equilibrium, as the cavities involved are heated to the temperature of operation. But given that all laboratory blackbodies suffer the same shortcomings, the production of super-Planckian emission in the near and far fields cannot be easily dismissed. After all, in order for Planck to obtain a blackbody spectrum in every arbitrary cavity, he had to drive the reflection term, either by injecting a carbon particle or by permitting additional heat to enter the system, beyond that required at the onset of thermal equilibrium"* [23]. Johnson cannot apply his arguments to metamaterial experiments and not make them with regard to regular laboratory cavities. In light of these many considerations, he has not demonstrated that my position, relative to the universality of blackbody radiation, has been overstated.

2.9 Robtaille's thought experiment

In the next section of his paper, Johnson reviews a very simple thought experiment, which I advanced in 2014, illustrating that Kirchhoff's Law cannot be valid [20]. Briefly, the idea involved two cavities. The larger outer cavity was constructed from perfectly emitting and absorbing walls and initially placed in a helium bath. Within this cavity and *in thermal contact with its floor*, rested an inner cavity made from perfectly reflecting walls. Initially, one of the six sides of this latter cavity remained open. As such, both cavities now contained black radiation at 4K, which had been produced by the outer cavity. The open wall of the inner cavity was then closed. It thus contained blackbody radiation at 4K. Then, the helium bath was removed and the system was allowed to rise to room temperature. In that case, the inner cavity still contained radiation associated with a 4K blackbody and the outer cavity contained radiation corresponding to room temperature.

Johnson argues that: *But by making the inner cavity walls perfectly reflecting and closing the last side, Robitaille has created two entirely separate cavities; by definition, the inner cavity walls cannot emit radiation in either direction, whatever the temperature. They therefore act as boundary walls to what has become a "hollow" outer cavity. The outer cavity no longer contains the inner cavity within itself in a thermal sense; Kirchhoff's Law therefore survives this thought experiment"* [36].

There is no validity in this argument. Simply examine the quotation by Kirchhoff which opens this reply: *"... In the interior therefore of an opake red-hot body of any temperature, the illumination is always the same, whatever be the constitution of the body in other respects"*. Obviously, Kirchhoff's statement has been violated and Kirchhoff's law permits the placement of any object within the cavity interior, provided that it does not have the ability to emit photons by non-thermal means. I have not sidestepped the conditions set forth by Kirchhoff. The inner cavity, having perfectly reflecting walls, is linked to the floor of the outer cavity through thermal conduction [20]. The inner cavity is not composed of an adiabatic wall which is unable to contain or transmit heat, as Planck used. Rather, it is made of a perfect reflector, best approached by a material such as silver: *"Since the inner cavity is perfectly reflecting, it will also be highly conducting, as good reflectors tend to be good conductors"* [20, p. 38]. Therefore, conductive heat transfer was allowed [23, 26]. Silver is known to be essentially a perfect reflector in the infrared ($\rho > 0.994$ [54]), as I previously mentioned in the work under question [20]. It also possesses one of the highest electrical conductivities and has a very reasonable thermal conductivity, on the order of 400 W m^{-1} K^{-1} [55]. Johnson cannot argue that: *"The outer cavity no longer contains the inner cavity within itself in a thermal sense."* [36].

Mathematically, adiabatic walls can act as perfect reflectors, but reflectors themselves are not mathematical walls. Silver reflectors can be characterized by temperature, precisely because, though they are ideally immune to capturing radiative energy, they are able to allow energy to enter or leave either through conduction or, when applicable, convection. Johnson will not deny that thermal conduction exists. Conversely, adiabatic walls cannot be characterized by any temperature, as they are fully immune to energy transfer by radiation, conduction, and convection.

In the case of a perfect reflector, all of the energy of the system can be trapped in its walls. In the case of the perfect absorber, Planck considered that all of the energy was contained in the radiation field. Yet, Planck still needed to allow his oscillators the opportunity to have some momentary interaction with radiative energy. Otherwise, no photons could be produced or absorbed. Similarly, the perfect reflector must be allowed to have some momentary interaction with conductive energy. Johnson can no more deny the presence of thermal conduction than he can deny the presence of thermal emission and absorption. Silver, an near perfect reflector in the infrared, still has access to conductive paths of heat transfer.

Johnson tries to dismiss this thought experiment [20] and with good reason. It constitutes strong evidence that Kirchhoff's Law could never have been correct. In fact, let us revisit this setting, as it also helps to dispel Planck's ill-conceived claims relative to the carbon particle acting as a catalyst.

First, note that the radiation contained within the inner cavity depends on its history prior to the cavity being closed [20]. It will contain whatever radiation was present within the outer perfectly absorbing cavity at that time. That is, it will be defined by the temperature of the outer cavity at closure (i.e. 4 K). The radiation within the inner cavity persists as Planck claims [4, § 51], but in a state which was well-defined by history, not just any arbitrary state.

If we place a carbon particle in the perfectly reflecting cavity and if this particle does not act to transform heat from the wall into the radiative field, but can only act as a catalyst, as Planck claimed [4, § 51] (relative to the existing radiation which initially corresponded to 4 K radiation [20]), the interior of the cavity could never become black. That is because the interior of the second cavity lacks sufficient energy in its 4 K photons to adopt the proper blackbody intensity for the new higher temperature of its walls, when the both cavities have been brought to room temperature [20]. The carbon particle, can never act to shift the Wien's peak to higher frequencies because Planck denies that it can contain any significant heat on its own [4, § 51]. The cavity, in this instance, could not contain black radiation at room temperature, without violation of the First Law of Thermodynamics. That is the central problem in Planck's notion that the carbon particle was merely a catalyst. In the example provided, Planck would stand in violation of the First Law, if he persisted in insisting that the carbon particle was not transforming the energy content of the walls and if he maintained his insistence that the

cavity became filled with blackbody radiation.

Should the carbon particle be characterized with a temperature, but interaction with the walls still prevented, then it could convert the radiation within the cavity to the proper Planckian distribution for the higher temperature. But this radiation will always remain gray, as the temperature of the carbon particle cannot be allowed to fall and since it has no access to other sources of heat. Once again, Planck is restricted by the First Law. The relative distributions of frequencies might become correct, but their intensity will always be too low.

It is only when the carbon particle is allowed to transform the thermal energy contained within the wall of the perfectly reflecting cavity that we obtain the correct answer and that the interior of the second cavity can become black [15]. That is why the carbon particle was never a catalyst. Planck ignores its ability to transform thermal energy contained within the walls. He was only concerned with the radiation field and this was a crucial error.

3 Robitaille and Crothers' 2015 paper

This section begins, once again, by claiming that there was a *volte-face* relative to my position on Stewart's mechanism. As noted previously, such claims are unwarranted. I have never supported *"Stewart's mechanism"* as providing a valid means of extending Kirchhoff's claims to all cavities made from arbitrary materials. Neither has Steve Crothers.

While defending Max Planck, Johnson has failed to recognize that there can be a substantial difference between 1) what Planck claims to have done, 2) what he actually did, and 3) what nature permits. For instance, when Planck denied the absorptivity of the surface layer and inserted only reflectivity, he made claims which were *demonstrably false* in the laboratory, relative to the nature of a blackbody surface. He inappropriately applied polarized light and Brewster's Law to secure his proof, when such an approach was disallowed based on the very definition of heat radiation. Finally, he concluded that his unnumbered equation at the end of section § 36 [4],

$$\frac{\mathbf{K}_\nu}{\mathbf{K}'_\nu} \cdot \frac{q^2}{q'^2} = \frac{1 - \rho'}{1 - \rho},$$

could be satisfied by all values of ρ and ρ'. Yet, when $\rho = 1$, this expression became undefined. As such, both Crothers and I maintain that Planck's "proof" [4] of Kirchhoff's Law remains fundamentally flawed and invalid. Planck has, therefore, been deprived of any justification in claiming universality. In his initial paper [3] and in the latter portion of his text [4], Planck correctly derived an expression for the blackbody function. But Planck can never state, based on § 35-37, that interiors of all cavities contain black radiation. This remains a serious crack in the armor of modern physics and Johnson's letter has not helped to rectify the problem.

3.1 The meaning of Planck's term "surface"

Within his classic text, Planck described how he has deviated from Kirchhoff's definition of a blackbody. For Johnson, Planck's new definition was permitted, whereas, in truth, it constituted a rejection of nature itself. As we have highlighted [25, p. 124], Planck stated within a footnote *"In defining a blackbody Kirchhoff also assumes that the absorption of incident rays takes place in a layer "infinitely thin". We do not include this in our definition"* [4, § 10]. This was not footnote material, as it constitutes a critical redefinition of the blackbody. In opposition to Kirchhoff, Planck decided to write: *"The creation of a heat ray is generally denoted by the word emission. According to the principle of the conservation of energy, emission always takes place at the expense of other forms of energy (heat, chemical or electric energy, etc.) and hence it follows that only material particles, not geometrical volumes or surfaces, can emit heat rays. It is true that for the sake of brevity we frequently speak of the surface of a body as radiating heat to the surroundings, but this form of expression does not imply that the surface actually emits heat rays. Strictly speaking, the surface of a body never emits rays, but rather it allows part of the rays coming from the interior to pass through. The other part is reflected inward and according as the fraction transmitted is larger or smaller the surface seems to emit more or less intense radiations"* [4, § 2].

Was Kirchhoff actually correct? Does the absorption of incident rays take place in a layer *"infinitely thin"* [4, § 10]? Or, did Planck more closely approximate nature: *"Strictly speaking, the surface of a body never emits rays, but rather it allows part of the rays coming from the interior to pass through"* [4, § 2]. Of course, if a surface, strictly speaking, cannot emit rays, it also cannot absorb rays.

The answer to this problem has been provided in the laboratory. If one considers the hexagonal planar structure of graphite and the reality that soot (or lampblack) has always played an important role relative to the creation of blackbodies (see references within [16, 17]), then the answer is readily apparent. For soot shares, in large measure, the hexagonal planar structure of graphite, although more breaks exist in the lattice. The surface of graphite or soot, is well represented by graphene [56, 57], as this alone constitutes the outer layer of a sheet of graphite.

Mak et al [58] speak of the absorption of graphene, *"Indeed, it was the strong absorption of single-layer graphene (with its absorbance of ~2.3%, ... that permitted the initial discovery of exfoliated monolayers by visual inspection under an optical microscope"*. The authors are referring to the work of Novoselov and Geim [56] (Nobel Prize, Physics, 2010). Moreover, even a single layer of graphene has been shown to be an absolutely phenomenal emitter, when driven by current [59].

Consequently, Planck's position that, *"Strictly speaking, the surface of a body never emits rays, but rather it allows*

part of the rays coming from the interior to pass through" [4, § 2], simply cannot be upheld. Laboratory evidence is firm on this point: some of the rays will begin to be absorbed even by the first mono-layer of atoms and less than 50 hexagonal planes of atoms should result in near complete absorption.

It is a fact that, even a single layer of graphene, the only structure which can be associated with the surface of a graphite blackbody, has powerful absorbance. Max Planck cannot be permitted to neglect this layer, when discussing blackbodies. Planck knew Kirchhoff's definition and chose to ignore it, even though he recognized that he could make the entire radiation within a cavity black, by introducing even the smallest of carbon particles [4, § 51]. Planck's error was in not allowing any absorption or emission at all, not in allowing that a single layer did not have 100% absorption. In this respect, Planck's statement was imprudent at the time and laboratory experiments have now demonstrated that, indeed, it was false. Johnson cannot correct this situation. Yet, as we shall see below, this was a critical step towards Planck's faulty derivation of Kirchhoff's Law. As for Kirchhoff, given his period in history, it is clear that his definition remains valid. For a single layer of atoms is as "infinitely thin" as nature can allow and 50 layers of atoms about as "infinitely thin" as a man could conceive in Kirchhoff's days. He could have no concept of the dimensions of atoms in 1860 [1, 2].

In continuing his letter, Johnson then attempts to justify Planck's insistence that the term *"bounding surface"* referred to a geometrical surface dividing two media, and that *"the material effects of emission and absorption take place within the adjoining media"* [36]. In this respect, we return to the question of what Planck has said and what he can be permitted to say.

In § 35 of his textbook [4], Planck outlined the notation relative to primed and unprimed superscripts: *"Let the specific intensity of radiation of frequency ν polarised in an arbitrary plane be $\boldsymbol{K}_\nu$ in the first substance ..., and $\boldsymbol{K}'_\nu$ in the second, and, in general let all quantities referring to the second substance be indicated by the addition of an accent"* [4]. Planck continued in § 43, *"The most adequate method of acquiring more detailed information as to the origin and the paths of the different rays of which the radiations I_1, I_2, I_3, ... I_n consist, is to pursue the opposite course and to inquire into the future fate of that pencil, which travels exactly in the opposite direction to the pencil I and which therefore comes from the first medium in the cone $d\Omega$ and falls on the surface element $d\sigma$ of the second medium"* [4, § 43]. Here, Planck clearly assigned to the surface element $d\sigma$, properties of the second medium.

Johnson argues that Planck's bounding surface did not have to absorb any light, citing Planck's claim, *"Thus only material particles can absorb heat rays, not elements of surfaces, although sometimes for the sake of brevity the expression absorbing surfaces is used"* [4, § 12]. But what Johnson fails to understand is that, should he argue along these lines, he would be brought to accept yet another truth from Robitaille and Crothers which I now state: Only material particles can reflect light! Thus, Planck cannot be allowed an imaginary surface which reflects light, while at the same time denying that this same surface can absorb or emit light.

The truth being that when Planck placed two materials together, the bounding element, $d\sigma$, must be characterized on one side by the reflectivity and aborptivity of the first material and on the other side, by the reflectivity and absorptivity of the second material. That is because, the elements in either of the materials are not properly characterized only by reflectivity. This is precisely why Crothers and I object to Planck's use of a bounding surface which does not *fully represent* the materials which it unites.

Planck is welcome to claim that he can place a hypothetical bounding surface between two materials which considers only transmission and reflection. As for Crothers and I, we continue to object. The bounding surface which Planck envisioned was completely detached from reality. The issue is not that Planck cannot place the geometric surface between two layers. That is self-evident. The issue is that Planck cannot detach this geometric bounding layer from the material properties of those substances which he claims it characterizes. It is impossible to extract only the reflectivity of a particle, assign it to a geometric bounding surface, and at the very same time, ignore the absorptivity of this same particle. Contrary to Johnson, our statement that *"Planck neglected the fact that real materials can possess finite and differing absorptivities"* [25, p. 127] is entirely appropriate and valid. Planck's own textbook provides additional insight: *"Whenever absorption takes place, the heat ray passing through the medium under consideration is weakened by a certain fraction of its intensity for every element of path traversed"* [4, § 12]. By necessity, the element contained within the bounding section is one of the elements in the path traversed. Planck cannot ignore its absorption, because its properties can only be related to the medium to which it is linked.

Johnson then attempts to counter our statement: *"Third, the simplest means of nullifying the proof leading to Planck's Eq. 42, is to use a perfect reflector as the second medium. In that case, a refractive wave could never enter the second medium and Planck's proof fails"* [25, p. 127]. In order to counter this argument, Johnson tries to make the bounding surface perfectly reflecting, but unfortunately, he is not allowed to adopt such an approach, as Planck's proof intrinsically depends on the transmissivity of this bounding surface. Johnson cannot make it a perfect reflector, as in doing so, he optically isolates the two media. Furthermore, Johnson has failed to notice what has been mentioned above; namely, if $\rho = 1$, then $(1 - \rho) = 0$ and Planck's equation, at the bottom of Planck's § 36 (see § 3 herein) becomes undefined. Planck needs this equation to be valid in order to obtain his Eq. 41, $q^2\mathbf{K}_\nu = q'^2\mathbf{K}'_\nu$. But after he obtains Eq. 40, $\rho=\rho'$, Planck must return to the equation he lists at the end of § 36 and this ex-

pression is not always true.

It also remains the case that Planck's entire proof of Kirchhoff's Law collapses, as Crothers and I correctly highlighted in our joint paper [25, p. 127], when we replaced the second medium with a perfect reflector.

At first, it was difficult for me to even understand why Johnson would have wanted to replace the geometric bounding surface with a perfectly reflective surface. The answer rests in his use of this quote from Max Planck: *"Since the equilibrium is nowise disturbed, if we think of the surface separating the two media as being replaced for an instant by an area entirely impermeable to heat radiation, the laws of the last paragraphs must hold for each of the two substances separately"* [4, § 35]. However, in that case, Planck was referring to the treatment he had just outlined when addressing a single medium. Note that Planck writes in § 32, *"that the total state of radiation of the medium is the same on the surface as in the interior.* Then in § 33, Planck writes, *"While the radiation that starts from a surface element and is directed towards the interior of the medium is in every respect equal to that emanating from an equally large parallel element of area in the interior, it nevertheless has a different history. That is to say, since the surface of the medium was assumed to be impermeable to heat, it is produced only by reflection at the surface of radiation coming from the interior"* [4, § 35]. Planck had assumed that the surface in this case was impermeable to heat because this was the only way he could treat the isolated medium near its surface.

However, when Planck moved to two media, he no longer used a boundary impermeable to heat, but assumed that the surface of each medium was *"smooth"* [4, § 36]. In § 9, Planck had defined a smooth surface as one which can partially reflect and transmit the incoming radiation [4, § 9]. Planck required transmission for his later proof of Kirchhoff's law in § 35 and § 36. This is an essential element, which Johnson failed to consider in stating that a perfectly reflecting boundary enabled $\rho = \rho'$. In that case, as mentioned above, the equation at the bottom of Planck's § 36 would become undefined. It is for this reason that Johnson cannot support Planck's position, by making the bounding surface a perfect reflector.

Robitaille and Crothers remain correct. Planck improperly treated absorption and reflection in his derivation. Furthermore, the use of a perfect reflector for the second medium was all that was needed to shatter Planck's proof of Kirchhoff's Law, as we have previously noted [25].

3.2 Absorption and transmission

This section of Johnson's letter begins by quoting from the paper by Robitaille and Crothers: *"With his words, Planck redefined the meaning of a blackbody. The step, once again, was vital to his derivation of Kirchhoff's Law, as he relied on transmissive arguments to arrive at its proof. Yet, blackbody radiation relates to opaque objects and this is the first indication that the proofs of Kirchhoff's Law must not be centered on arguments which rely upon transmission. Planck ignored that real surface elements must possess absorption, in apparent contrast with Kirchhoff and without any experimental justification"* [25, p. 124].

Strangely, Johnson then concludes from this quotation that *"the apparent problem arises from the fact that Planck's surface is a geometrical one, whilst Robitaille and Crothers are obviously referring to a surface layer in which, they maintain, all absorption must take place because transmission is not permitted through a black body"* [36]. But we never stated that *all* of the absorption must take place from the surface layer. We stated that *"real surface elements must possess absorption"* [25, p. 124]. The surface need not have 100% absorption, as only a slight absorption is sufficient to invalidate Planck's proof. It is obvious, from our treatment of the first section of Planck's proof, that we do in fact allow transmission to take place within the medium and for elements within the blackbody to absorb, exactly like Max Planck [25, § 4.2]. We caution, however, that blackbodies are opaque objects and that Planck's proof cannot rely exclusively on transmission and reflection. Our point remains valid, as well demonstrated by the experimental realities outlined relative to graphene in § 3.1 above.

Again quoting from our paper, Johnson then attempts to argue that Planck was correct in inferring that *"... while in the case of bodies with vanishingly small absorbing power only a layer of infinite thickness may be regarded as black"* [4, § 10]. Once again, it is difficult to understand how Johnson can come to Planck's defense in this case. An opaque object which has a low absorptivity, also has a high reflectivity *by definition*. If not, it would not be opaque. As such, most photons which approach an opaque surface with low absorptivity are reflected away from the body. For Planck's argument to work, one would have to discount the surface reflection from an opaque object with a low emissivity which is counter to all laboratory experience. This highlights that Planck's new definition of a blackbody is completely outside the laws of nature. Planck cannot argue that he can neglect surface reflection, simply to salvage his derivation of Kirchhoff's Law.

Our point remains valid *"Blackbodies are opaque objects without transmission, by definition"* [25, p. 125]. Still, we have, in fact, allowed Planck to have some mathematical latitude and some level of transmission within the object, as presented in our § 4.2 [25]. But we cannot allow Planck to completely negate the presence of the reflection which is known to occur at the surface of an opaque object of low emissivity. Johnson and Planck shall not redefine nature.

3.3 Reflection

Relative to neglecting the reflection which occurs within a medium, we never stated that such an approach was invalid,

merely that it was suboptimal. In fact, in § 4.2 of our paper, we specifically outline the effects of neglecting the reflection taking place within the medium [25].

Johnson, however, is under the impression that reflection is strictly a surface phenomenon and cannot take place within the medium. At the end of this section, Johnson emphasizes the point when he states *"Note that Planck is still talking about the interior of the medium where reflection is not applicable because there is no surface; therefore Robitaille and Crothers' objection cannot be maintained"* [36, § 3.3]. Johnson is confused on this point.

Planck himself explicitly commented on scattering within media: *"The propagation of the radiation in the medium assumed to be homogeneous, isotropic, and at rest takes place in straight lines and with the same velocity in all directions, diffraction phenomena being entirely excluded. Yet, in general, each ray suffers during its propagation a certain weakening, because a certain fraction of its energy is continuously deviated from its original direction and scattered in all directions. This phenomenon of "scattering", ... takes place, generally speaking, in all media differing from absolute vacuum ... "* [4, § 8]. Later in the same section, Planck noted that, beyond diffraction, scattering also depends on reflection [4, § 8]. Hence, contrary to Johnson's claims, Planck understood that reflection is not strictly a surface phenomenon.

Crothers and I have properly considered internal reflection [25, § 4.2]. We have demonstrated that, when internal reflection is considered, powerful new insight is gained. Rather than simply obtaining Kirchhoff's formulation, $\mathbf{K}_\nu = \boldsymbol{\epsilon}_\nu/\alpha_\nu$, which is potentially undefined, we can actually extract $\boldsymbol{\epsilon}_\nu = (1-\rho_\nu)\mathbf{K}_\nu$, which is never undefined [25, § 4.2]. The insight provided by this treatment is important, contrary to what Johnson implies when insisting, without justification and in opposition to Planck's own statements, that reflection is only a surface phenomenon.

3.4 Polarization and equality of reflection

In the final section of his letter, Johnson attempts to justify Planck's use of polarized light and his assertion that the reflectivities of a pair of media at the bounding surface must be equal. He begins by quoting from our paper: *"In § 5 Planck admitted that homogeneous isotropic media emit only natural or normal, i.e. unpolarized, radiation: "Since the medium was assumed to be isotropic the emitted rays are unpolarized". This statement alone, was sufficient to counter all of the arguments which Planck later utilized to arrive at Kirchhoff's Law [Eq. 42]. That is because the important sections of Planck's derivation, namely § 35–37 make use of plane-polarized light. These steps were detached from experimental reality, relative to heat radiation [Planck, § 35] ... "*.

At this point, Johnson recalls that we have allowed Planck to resolve heat radiation into two equal orthogonal components, each plane-polarized. He objects to our statement that *"such rays could never exist in the context of heat radiation"* [25, p. 129]. Apparently, Johnson has failed to grasp that even though Planck can resolve heat radiation into two components, he is not allowed to apply only one component in his derivation. He must always consider *both* components, even if he can resolve them into two orthogonal planes.

Johnson apparently does not understand why Planck wanted to treat only one component, in part, because he seems unaware of Brewster's Law. Planck, in his derivation of Kirchhoff's Law, invoked plane-polarized radiation, such that he could set $\rho = \rho' = 0$. He could only obtain this expression, when dealing with a single plane polarized beam of light. That is because, if he sent such a beam at the proper angle and with the proper polarization towards his bounding surface, there would be no reflection, according to Brewster's Law.

However, Planck was not right in stating that there could be no reflection in the context of heat radiation. He could not obtain the plane-polarized beam of light, which he required, because the other component of the radiation, which was inappropriately ignored in his derivation, was also present. Moreover, Planck did not even test reflectivity by his argument from Brewster's law, as the latter is dependent upon the presence of a reflected ray as well as a transmitted ray. Thus, Planck could not conclude that the reflectivities of both materials were 0. The absence of a reflected ray does not imply that reflectivity is zero, as the polariscope attests. Just because Planck can resolve light into two components does not mean that he can ignore one of these components. *This is one of the most significant flaws in Planck's derivation of Kirchhoff's Law*.

Johnson then tries to defend Planck's most dramatic claim. Planck states [4, § 37]: *"Now in the special case when the rays are polarized at right angles to the plane of incidence and strike the bounding surface at the angle of polarization,* $\rho = 0$, *and* $\rho' = 0$. *The expression on the right side of the last equation then becomes 1; hence it must always be 1 and we have the general relations:*

$$\rho = \rho' \qquad (40)$$

and

$$q^2\mathbf{K}_\nu = q'^2\mathbf{K}'_\nu \quad (41)".$$

As I have just outlined, Planck cannot refer to this special case, because he does not have access to light polarized in a single plane. He must *always simultaneously treat both components*. Secondly, Planck is incorrect in asserting that the right side of the expression at the end of his § 36 [4] (also shown in § 3 of this letter), *"must always be 1"*, because it becomes undefined when $\rho = 1$. Planck was making an elementary error in mathematics. We maintain that *"The result was stunning."* [25, p. 129]. We also maintain that *"Max Planck had determined that the reflectivities of all arbitrary media were equal"* [25, p. 129].

Johnson then tries to defend Planck one last time, by insisting that what the latter *"had in fact demonstrated is that the reflectivities on each side of a geometrical surface bounding two different media are equal. Clearly if a different pair of media are chosen, the value of the reflectivity of the bounding surface may be different as well"* [36, § 3.4]. He then quoted from *The Theory of Heat Radiation*, *"Since, in general, the properties of a surface depend on both of the bodies which are in contact, this condition shows that the property of blackness as applied to a body depends not only on the nature of the body but also on that of the contiguous medium. A body which is black relatively to air need not be so relatively to glass, and vice versa"* [4, § 10].

Both Crothers and I understand what Max Planck claimed. However, we are properly concerned with what he has actually done. Planck's statement that *"the properties of a surface depend on both of the bodies which are in contact"* [4, § 10] can never be verified in the context of opaque media, precisely because his bounding surface is an abstraction. Snell's law, for instance, also relies on the interface of two media, but a bounding surface, or the changes at the surfaces, need not be introduced to obtain the proper answer. The indices of refraction of the two media alone are sufficient to treat the problem.

Planck's statements relative to the bounding surface were subject to two fundamental objections. First, they are justified by nothing; second, they constitute *"une hypothèse gratuite"* (see table presenting arguments against 19th century proofs of Kirchhoff's Law in [47, p. 16]). Planck may wish to claim that *"A body which is black relatively to air need not be so relatively to glass, and vice versa"* [4, § 10], but he had absolutely no justification for such a statement.

Rather, what Planck did possess are two isotropic media. Each of these is characterized by the absorptivity and reflectivity for each of its constitutive elements. Within his bounding surface, Planck could only introduce the reflectivity of elements contained in the media in question. When he introduced this reflectivity into his bounding surface, he had to additionally introduce some absorptivity, since this also characterized the media. Planck was not free to ignore the absorptivity. But he did so, as absorptivity in the bounding surface would prevent him from making use of Brewster's Law.

In any case, Planck could not invent a new reflectivity, which now existed only when he places the two media in contact with one another. After all, the reflectivities of the bounding surface must somehow be related to the materials under study. Furthermore, all that Planck could ever know about these materials are the reflectivities which can be measured. Neither he, nor Johnson, are allowed to hypothesize on what can never be measured in opaque media.

Planck recognized that he could not state that reflectivities of all materials are identical. As such, he postulated, *without any experimental evidence*, that his proof actually refers to something else [4, § 10]. Crothers and I dispute such claims. Planck's derivation must be taken on what the setting and the mathematics demonstrate. If we ignored Planck's mathematical errors and experimental oversights, we could much more convincingly argue that he had demonstrated that the reflectivities of all arbitrary materials were equal, using the same proof. Planck could measure nothing more than the reflectivities of each medium. Thus, he remains in violation of known optics, despite his attempts to introduce a new meaning to the reflectivity of a surface. Furthermore, Planck is forbidden from writing Eq. 40, $\rho = \rho'$, precisely because he has violated nature's rule that heat radiation is never polarized. It also remains the case that the unnumbered equation, which Planck presents at the end of his § 36 [4] (see § 3 herein), is undefined when $\rho = 1$.

4 Johnson's summary and conclusions

In opening this section of his letter, Johnson claims that, *"Stewart [33] had shown that the radiation in a cavity made from perfectly absorbing material at thermal equilibrium must be black, of an intensity appropriate to the equilibrium temperature. According to Robitaille, Kirchhoff [1] extended this finding to cavities made of arbitrary materials"*. Once again, Johnson has missed the mark.

Stewart considered plates in his experiments and Johnson is distorting what Stewart has done. It was with plates that Stewart demonstrated the Law of Equivalence (in modern notation: $\epsilon = \alpha$, or $\epsilon + \rho = \alpha + \rho$). Kirchhoff's extension to all arbitrary cavities [1, 2], went well beyond Stewart's legitimate law and has never been demonstrated to be true in the laboratory.

Johnson's claim that I have now withdrawn my objections to *"Stewart's mechanism"*, in my paper with Crothers [25], is without basis. *"Stewart's mechanism"* has numerous problems, including potential violations of the First Law of Thermodynamics, depending on the circumstances considered. It suffers from the reality that cavities made of low emissivity materials can prefer to increase the temperature of the walls, rather than emit a photon. Johnson's letter does nothing to counter this argument and that is why *"Stewart's mechanism"* cannot be realized in practice, as recognized by Crothers and myself.

Finally, Johnson admits: *"Robitaille is obviously correct to point out that black body cavities are never made from reflective materials."*. However, he then attempts to excuse the observation, in noting that, *"... this fact appears to be more a question of practicality and the need to ensure that the walls are not perfectly reflective at any wavelength so that proper measurements of temperature can be made. It does not seem to amount to a demonstration that Kirchhoff's Law necessarily fails, as Robitaille claims."*. Again, the arguments are ill-conceived. The fact that an experiment, required to establish a law of physics, still remains impractical after 150 years, well indicates that the law was never valid.

Some have argued, for instance, that when cavities are constructed from low emissivity materials, their dimensions need to be increased. This helps to augment their absorbance when sampling return losses. But Kirchhoff's law is explicit. The dimensions of the walls are irrelevant and, at a given temperature, must remain unrelated to the emissivity of the material, provided that the diffraction limit is avoided. The diffraction limit is not set by the emissivity. Furthermore, Johnson's arguments, relative to the ability to properly measure a temperature remain unfounded. Temperature sensors in the walls of cavities can easily report such information.

5 Conclusions

Throughout his letter, Johnson demonstrates that he has not carefully considered what Stewart, Kirchhoff, and Planck have written. He attributes to them positions which they never adopted. Then, he misinterprets the positions they did take. He repeatedly makes elementary errors relative to the understanding of cavity radiation. His statements on properly measuring the temperature of a cavity are but one example. He argues for *"Stewart's mechanism"*, in building up the radiation within a cavity, while not recognizing that the introduction of specular reflection within such objects can easily lead to the formation of standing waves. He also fails to understand that a cavity can simply increase the temperature of its walls and not emit a single photon.

He rejects my experimental work on MRI cavities, as unrelated to the problem of thermal emission and notes that processes, like fluorescence, have been excluded by Max Planck. Yet, the sampling of a cavity with a network analyzer does not involve such processes. In this respect, he also fails to note that Kirchhoff had mistakenly included such processes, in his initial work [1]. This was the only work of Kirchhoff which Johnson cited.

Furthermore, he fails to recognize that microwave cavities are utilized aboard modern satellites, wherein such objects are claimed to be black. Johnson also improperly and unknowingly expresses Stewart's Law as $\epsilon_\nu = \alpha_\nu + \rho_\nu$ in a thought experiment, thereby reaching conclusions which were clearly false. Then, he ignores the very existence of thermal conduction, when he attempts to invalidate my thought experiment with two cavities. He misrepresents my statements and those of Stephen Crothers, when he tries to state that we denied that the interior of a medium can have absorbance. He failed to understand the difference between resolving a heat ray into its two plane-polarized components and making use of a single plane-polarized ray, in order to infer something about heat radiation, which is never polarized. He hypothesized that replacing Planck's geometrical bounding surface with a perfect reflector could be used to validate Planck's claims, when clearly, it leads to an undefined mathematical expression and an invalid setting.

For all these reasons, Johnson cannot state that he has, in any way, nullified my objections to Kirchhoff's Law. Still, he must not be faulted for trying to defend Kirchhoff and Planck. As I stated in the introduction, it is the first obligation of a scientist to defend established science. Moreover, the study of cavity radiation is not at all simple. In this regard, Johnson's efforts are noteworthy and he is to be given credit for the time he has invested in reviewing these many papers.

Through the exchange prompted by his letter, Johnson has been indirectly responsible for bringing to the forefront many aspects of cavity radiation. Progress is often achieved, only when old ideas are first rejected, even if the process of discovery is not smooth. The process of correction, in itself, leads to scientific advancement. Hence, through such an exchange, readers can better come to understand why Kirhchoff's Law of thermal emission was never valid. Consequently, Planck's claims for universality must be rejected.

Acknowledgments

The author would like to thank the editors of *Progress in Physics*, namely, Dmitri Rabounski, Florentin Smarandache, and Larissa Borrisova, for the courage they have shown in publishing my work on Kirchhoff's Law, the Sun, and the Microwave Background. I am well aware that this has not been always easy for the Journal. I would also like to thank my colleague, Stephen J. Crothers, for valuable discussions relative to this work.*

Dedication

This work is dedicated to my youngest sister, Mireille and her husband, John.

Submitted on: January 25, 2016 / Accepted on: January 27, 2016
First published online on: January 28, 2016

References

1. Kirchhoff G. Über das Verhältnis zwischen dem Emissionsvermögen und dem Absorptionsvermogen. der Körper fur Wärme und Licht. *Poggendorfs Annalen der Physik und Chemie*, 1860, v. 109, 275–301. (English translation by F. Guthrie: Kirchhoff G. On the relation between the radiating and the absorbing powers of different bodies for light and heat. *Phil. Mag.*, 1860, ser. 4, v. 20, 1–21).
2. Kirchhoff G. Über den Zusammenhang zwischen Emission und Absorption von Licht und. Wärme. *Monatsberichte der Akademie der Wissenschaften zu Berlin*, sessions of Dec. 1859, 1860, 783–787.
3. Planck M. Über das Gesetz der Energieverteilung im Normalspektrum. *Annalen der Physik*, 1901, v. 4, 553–563 (English translation by ter Haar D.: Planck M. On the theory of the energy distribution law in the normal spectrum. The old quantum theory. Pergamon Press, 1967, 82–90; also Planck's December 14, 1900 lecture *Zur Theorie des Gesetzes der Energieverteilung in Normalspectrum*, which stems from this paper, can be found in either German, or English, in: Kangro H. Classic papers in physics: Planck's original papers in quantum physics. Taylor & Francis, London, 1972, 6–14 or 38–45).
4. Planck M. The Theory of Heat Radiation. P. Blakiston' Son & Co., Philadelphia, PA, 1914.

*Johnson and I had indeed exchanged correspondence following our meeting at a conference in 2014 [60].

5. Eddington A.S. The Internal Constitution of the Stars, Cambridge University Press, Cambridge, 1926.
6. Chandrashekhar S. An Introduction to the Study of Stellar Structure, Chicago University Press, Chicago, 1939.
7. Kippenhahn R. and Weigert A. Stellar Structure and Evolution, Springer-Verlag, New York, 1990.
8. Srinivasan G. From White Dwarfs to Black Holes: The Legacy of S. Chandrasekhar, The University of Chicago Press, Chicago, 1996.
9. Penzias A.A. and Wilson R.W. A measurement of excess antenna temperature at 4080 Mc/s. *Astrophys. J.*, 1965, v. 1, 419–421.
10. Dicke R.H., Peebles P.J.E., Roll P.G., and Wilkinson D.T. Cosmic black-body radiation. *Astrophys. J.*, 1965, v. 1, 414–419.
11. Robitaille P.M. Liquid metallic hydrogen: A building block for the liquid Sun. *Prog. Phys.*, 2011, v. 3, 69–74.
12. Robitaille P.M. Forty lines of evidence for condensed matter — The Sun on trial: Liquid metallic hydrogen as a solar building block. *Prog. Phys.*, 2013, v. 4, 90–142.
13. Robitaille P.-M. On the validity of Kirchhoff's law of thermal emission. *IEEE Trans. Plasma Sci.*, 2003, v. 31, no. 6, 1263–1267.
14. Robitaille P.-M. An analysis of universality in blackbody radiation. *Prog. Phys.*, 2006, v. 2, 21-23.
15. Robitaille P.-M. A critical analysis of universality and Kirchhoff's law: A return to Stewart's law of thermal emission. *Prog. Phys.*, 2008, v. 3, 30–35.
16. Robitaille P.-M. Blackbody radiation and the carbon particle. *Prog. Phys.*, 2008, v. 3, 36–55.
17. Robitaille P.-M. Kirchhoff's law of thermal emission: 150 Years. *Prog. Phys.*, 2009, v. 4, 3–13.
18. Robitaille P.-M. Blackbody radiation and the loss of universality: Implications for Planck's formulation and Boltzmann's constant. *Prog. Phys.*, 2009, v. 4, 14–16.
19. Robitaille P.-M. A radically different point of view on the CMB in *"Questions of Modern Cosmology"*, Mauro D'Onofrio and Carlo Burigana, eds., Springer Scientific, New York, N.Y., 2009, 93–108.
20. Robitaille P.-M. Further Insight Relative to Cavity Radiation: A Though Experiment Refuting Kirchhoff's Law. *Prog. Phys.*, 2014, v. 10, no. 1, 38–40.
21. Robitaille P.-M. Further Insight Relative to Cavity Radiation II: Gedanken Experiments and Kirchhoff's Law. *Prog. Phys.*, 2014, v. 10, no. 2, 116–120.
22. Robitaille P.-M. On the Equation which Governs Cavity Radiation I. *Prog. Phys.*, 2014, v. 10, no. 2, 126–127 (see also, Robitaille P.M. Notice of Revision: "On the Equation which Governs Cavity Radiation I, II", by Pierre-Marie Robitaille. *Prog. Phys.*, 2015, v. 11, no. 1, 88).
23. Robitaille P.-M. On the Equation which Governs Cavity Radiation II. *Prog. Phys.*, 2014, v. 10, no. 3, 157–162 (see also, Robitaille P.M. Notice of Revision: "On the Equation which Governs Cavity Radiation I, II", by Pierre-Marie Robitaille. *Prog. Phys.*, 2015, v. 11, no. 1, 88.).
24. Robitaille P.M. Blackbody radiation in optically thick gases. *Prog. Phys.*, 2014, v. 10, no. 3, 166–168.
25. Robitaille P.-M. and Crothers S.J. "The Theory of Heat Radiation" Revisited: A Commentary on the Validity of Kirchhoff's Law of Thermal Emission and Max Planck's Claim of Universality. *Prog. Phys.*, 2015, v. 11, no. 2, 120–132.
26. Robitaille P.-M. Further Insight Relative to Cavity Radiation III: Gedanken Experiments, Irreversibility, and Kirchhoff's Law. *Prog. Phys.*, 2016, v. 12, no. 1, 85–88.
27. Robitaille P.-M.L., Abduljalil A.M., Kangarlu A., Zhang X., Yu Y., Burgess R., Bair S., Noa P., Yang L., Zhu H., Palmer B., Jiang Z., Chakeres D.M., and Spigos D. Human magnetic resonance imaging at eight Tesla. *NMR Biomed.*, 1998, v. 11, 263–265.
28. Robitaille P.-M.L. Magnetic resonance imaging and spectroscopy at very high fields: a step towards 8 Tesla, in "Proceedings of Physical Phenonema at High Magnetic Fields-III", (Z. Fisk, L. Gor'kov, D. Meltzer and R.Schrieffer), eds., World Scientific, London, 1999, 421–426.
29. Robitaille P.-M.L. and Berliner L.J. (Eds.) Biological Magnetic Resonance. Volume 26: Ultra High Field Magnetic Resonance Imaging. Springer, New York, 2006.
30. Bloch F. Nuclear Induction. *Phys. Rev.*, 1946, v. 70, no. 7–8, 460–474.
31. Wien W. Über die Energieverteilung in Emissionspektrum eines schwarzen Körpers. *Ann. Phys.*, 1896, v. 58, 662–669.
32. Stefan J. Über die Beziehung zwischen der Warmestrahlung und der Temperature. *Sitzungsberichte der mathematischnatur wissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften*, Wien 1879, v. 79, 391–428.
33. Stewart B. An account of some experiments on radiant heat, involving an extension of Prévost's theory of exchanges. *Trans. Royal Soc. Edinburgh*, 1858, v. 22, no. 1, 1–20 (also found in Harper's Scientific Memoirs, edited by J.S. Ames: The Laws of Radiation and Absorption: Memoirs of Prévost, Stewart, Kirchhoff, and Kirchhoff and Bunsen, translated and edited by D.B. Brace, American Book Company, New York, 1901, 21–50).
34. Robitaille P.M. Stellar opacity: The Achilles heel of the gaseous Sun. *Prog. Phys.*, 2011, v. 3, 93–99.
35. Robitaille P.M. Water, hydrogen bonding, and the microwave background. *Prog. Phys.*, 2009, v. 2, L5–L8.
36. Johnson R.J. A Re-examination of Kirchhoff's Law of Thermal Radiation in Relation to Recent Criticisms. *Prog. Phys.*, 2016, v. 12, no. 3, 175–183.
37. Siegel D.M. Balfour Stewart and Gustav Robert Kirchhoff: Two independent approaches to Kirchhoff's Law. *Isis*, 1976, v. 67, no. 4, 565–600.
38. Hoffmann D. On the experimental context of Planck's foundation of quantum theory. In: Revisiting the Quantum Discontinuity, Max Planck Institute for the History of Science, Preprint 150, 2000, 47–68.
39. Lummer O. and Pringsheim E. Kritisches zur schwarzen Strahlung. *Annalen der Physik*, 1901, v. 6, 192–210.
40. Rubens H. and Kurlbaum F. Anwendung der Methode der Reststrahlenzur Prüfung der Strahlungsgesetzes. *Annalen der Physik*, 1901, v. 2, 649–666; Rubens H. and Kurlbaum F. On the heat radiation of longwave-length emitted by black bodies at different temperatures. *Astrophys. J.*, 1901, v. 74, 335–348.
41. Touloukian Y.S., DeWitt D.P. Thermal radiative properties of nonmetallic solids. Vol. 8 in: Thermophysical Properties of Matter, IFI/Plenum, New York, 1972.
42. Harris L. The optical properties of metal blacks and carbon blacks. MIT and The Eppley Foundation for Research, Monograph Ser., 1, New Port, R.I., Dec. 1967.
43. Nelms N., Dowson J. Goldblack coating for thermal infrared detectors. *Sensors and Actuators*, 2005, v. A120, 403–407.
44. O'Neill P., Ignatiev A, Doland C. The dependence of optical properties on the structural composition of solar absorbers: Gold black. *Solar Energy*, 1978, v. 21, 465–478.
45. Harris L., McGuiness R.T., Siegel B.M. The preparation and optical properties of gold black. *J. Opt. Soc. Am.*, 1948, v. 38, 582.
46. Kirchhoff G.R. On the relation between the emissive and the absorptive power of bodies for heat and light, In "Investigations on the Solar Spectrum and the Spectra of the Chemical Elements", 2nd Edition, Ferd. Dummler's Publishing House, Berlin, 1866; Gesammelte Abhandlugen, pp. 571-598, Leipzig, 1882; also found in Harper's Scientific Memoirs, edited by J.S. Ames: The Laws of Radiation and Absorption: Memoirs of Prévost, Stewart, Kirchhoff, and Kirchhoff and

Bunsen, translated and edited by D.B. Brace, American Book Company, New York, 1901, 73–97).

47. Schirrmacher A. Experimenting theory: The proofs of Kirchhoff's radiation law before and after Planck. *Hist. Stud. Phys. Biol. Sci.*, 2003, v. 33, 299–335.
48. Robitaille P.-M. Calibration of microwave reference blackbodies and targets for use in satellite observations: An analysis of errors in theoretical outlooks and testing procedures. *Prog. Phys.*, 2010, v. 3, 3–10.
49. Robitaille P.-M. The Planck Satellite LFI and the Microwave Background: Importance of the 4 K Reference Targets. *Prog. Phys.*, 2010, v. 3, 11–18.
50. Robitaille P.M. COBE: A radiological analysis. *Prog. Phys.*, 2009, v. 9, no. 4, 17-42.
51. Guo Y., Cortez C.L., Molesky S., and Jacob Z. Broadband super-Planckian thermal emission from hyperbolic metamaterials. [**?**], 2012, v. 101, 131106.
52. Yu Z., Sergeant N.P., Skauli T., Zhang G., Wang H., and Fan S. Enhancing far-field thermal emission with thermal extraction. *Nature Comm.*, 2013, DOI: 10.1038/ncomms2765.
53. Yu Z., Sergeant N., Skauli T., Zhang G., Wang H. and Fan S. Thermal extraction: Enhancing thermal emission of finite size macroscopic blackbody to far-field vacuum. 4 Nov 2012, arXiv:1211.0653v1 [physics.optics].
54. Diggers R.G (Ed), Encyclopedia of Optical Engineering (vol. 3). Marcel Dekker, Inc., New York, 2003, p. 2354.
55. Electrical Resistivity of Pure Metals. In, CRC Handbook of Chemistry and Physics, 2013–2014, CRC Press, Baca Raton, FL, p. 12–42.
56. Novoselov K.S., Geim A.K., Morozov S.V., Jiang D., Zhang Y., Dubonos S.V., Grigorieva I.V., Firsov A.A., Electric field effect in atomically thin carbon films. *Science*, 2004, v. 306, 666–669.
57. Geim A.K. and Novoselov K.S. The rise of graphene. *Nature Materials*, 2007, v. 6, 183–191.
58. Mak K.F., Ju L., Wang F., and Heinz T.F. Optical spectroscopy of graphene: From the far infrared to the ultraviolet. *Solid State Comm.*, 2012, v. 152, 1341–1349.
59. Kim Y.D., Kim H., Cho Y., Ryoo J.H., C.-H. Park, Kim P., Kim Y.S., Lee S., Li Y., Park S.-N., Yoo Y.S., Yoon D., Dorgan V.E., Pop E., Heinz T.F., Hone J., Chun S.-H., Cheong H., Lee S.W., Bae M.-H., and Park Y.D. Bright visible light emission from graphene. *Nature Nano.*, 2015, v. 10, 676–682.
60. Johnson Robert. "Could Electricity Explain Some'Known Unknowns' in Geology?" and "How Do Tall Trees Survive? New Models for the Flow of Sap", presented at The Electric Universe Conference, Albuquerque, N.M., March 20–24, 2014.

LETTERS TO PROGRESS IN PHYSICS

Errata to "Mansouri-Sexl Test Theory: The Question of Equivalence between Special Relativity and Ether Theories"

Maciej Rybicki
E-mail: maciej.rybicki@icloud.com

The title paper [1] contains an essential mistake committed by the present author. Namely, the Mansouri and Sexl generalized transformation of time, as well as the relevant form of the Lorentz transformation of time have been erroneously read and typed. In consequence, a certain part of the paper (indicated in the table, below) requires replacing to conform to the correct equations. The rest of the paper, except for the minor corrections indicated in the Errata, still remains valid. Let me apologize to the Readers and Editors for the inconvenience.

Page	Written	Read
89 (Abstract)	"has an erroneous form."	"is incorrectly used."
89, Eq. (2)	$t = aT + \epsilon X$	$t = aT + \epsilon x$
90, Eq. (5)	$x = \gamma(X - vt)$	$x = \gamma(X - vT)$
91 (Conclusion)	"We have shown that an incorrect notation…"	"We have shown that an incorrect use…"

Page 90, left column

Written (part to be replaced, starting from):
"Mansouri and Sexl state that for $a = b = 1$, $\epsilon = 0$ the Galilean transformation is obtained, which is correct. …"
(ending with, 33 lines down):
"…Consequently, they concluded that only violation of the two-way isotropy resulting in deviations from the relativistic values of a and b constitutes a challenge to STR."

Read (part to be introduced):
"Thus, the difference in the one-way speed of light would be a sole matter of choice of the synchronization convention.

From M-S theory it follows that for $a = b = 1$, $\epsilon = 0$, the Galilean transformation is obtained. If, after employing the external synchronization, a and b equal to unity, it would mean that mechanical phenomena are ruled by Newtonian physics and subject to the Galilean principle of relativity, while the Maxwell equations (and the relevant constant speed of light) refer to the ether frame only.

Instead, for $1/a = \gamma$ and $\epsilon = -v/c^2$, the M-S transformation of time turns into the Lorentz transformation of time:

$$t' = \frac{t}{\gamma} - \frac{vx'}{c^2}\,. \tag{E1}$$

In this form, the "rest-to-observer" coordinates appear on both sides of equation. Written in the same manner, the inverse Lorentz transformation is therefore:

$$t = \frac{t'}{\gamma} + \frac{vx}{c^2}\,. \tag{E2}$$

Consequently, the M-S transformation of time, and the inverse transformation are:

$$\begin{aligned} t &= aT + \epsilon x\,,\\ T &= at - \epsilon X\,. \end{aligned} \tag{E3}$$

Now, assuming $1/a = \gamma$ and $\epsilon = 0$, we obtain:

$$t = \frac{T}{\gamma} \Longrightarrow T = t\gamma\,, \tag{E4}$$

in contradiction with

$$T = \frac{t}{\gamma}\,. \tag{E5}$$

Mansouri and Sexl intended to treat independently the questions of time dilation and simultaneity. This, however, is infeasible with respect to the Lorentz transformation in which relativity of simultaneity and relativistic effects are inseparably connected. In the Lorentz transformation, one cannot obtain time dilation without taking into account the relativity of simultaneity. Likewise, the self-consistence of reciprocal equations in the Lorentz transformation involves the mutual dependence between $\gamma = 1/\sqrt{1 - v^2/c^2}$ and v/c^2. The incorrect use of Lorentz transformation (in particular, not including the inverse transformation) led to a false conclusion as to the question of equivalence between STR and the postulated ether theory."

Submitted on January 30, 2016 / Accepted on February 10, 2016

References

1. Rybicki M. Mansouri-Sexl Test Theory: The Question of Equivalence between Special Relativity and Ether Theories. *Progress in Physics*, 2016, v. 12 (1), 89–92.

Application of the Differential Transform Method to the Advection-Diffusion Equation in Three-Dimensions

Samia H. Esmail and Mahmoud M. Taha

Mathematics and Theoretical Physics Department, Nuclear Research Center, Atomic Energy Authority, Cairo, P.No. 13759, Egypt.
E-mail: mahmoudmt@hotmail.com

Advection diffusion equation with constant and variable coefficient has a wide range of practical, industrial and environmental applications. Due to the importance of atmospheric dispersion equation, we present this study which deals analytically with the atmospheric dispersion equation. The present model is proposed to estimate the concentration of an air pollutant in an urban area. The model is based on using Differential Transform Method (DTM) to solve the atmospheric dispersion equation. The model assumes 1) the pollutant is released from an elevated continuous point source; 2) there exist an elevated inversion layer; 3) the dispersion coefficients are parameterized as a function of downwind distance in a power law dependence. To test the model accuracy, the model predictions have been applied and compared with the experimental data for the Inshas research reactor (Egypt). The model predictions are shown to be in good agreement with the measurement of field data.

1 Introduction

The advection-diffusion equation of air pollution in the atmosphere is essentially a statement of conservation of the suspended material. The concentration of turbulent fluxes are assumed to be proportional to the mean concentration gradient which is known as Fick-theory.

This assumption, combined with the continuity equation, leads to the steady-state advection-diffusion equation, Blackadar [1]

$$\frac{\partial C}{\partial t}+u\frac{\partial C}{\partial x}+v\frac{\partial C}{\partial y}+w\frac{\partial C}{\partial z}=\frac{\partial}{\partial x}\left(k_x\frac{\partial C}{\partial x}\right)+\\+\frac{\partial}{\partial y}\left(k_y\frac{\partial C}{\partial y}\right)+\frac{\partial}{\partial z}\left(k_z\frac{\partial C}{\partial z}\right) \tag{1}$$

where $C(x,y,z)$ denotes the concentration, k_x, k_y, k_z are the cartesian components of eddy diffusivity and u, v, w are the cartesian components of wind speed, where x, y are cartesian horizontal distance and z is the height above ground surface.

In order to solve (1) we included the following assumptions: the pollutants are inert and have no additional sinks or sources downwind from the point source, the vertical w and lateral v components of the mean flow are assumed to be zero, k_x is neglected, k_y and k_z are functions of downwind distance. The mean horizontal flow is incompressible and horizontally homogeneous (steady state). Then, (1) is simplified to be:

$$u\frac{\partial C}{\partial x}=k_y\left(\frac{\partial^2 C}{\partial y^2}\right)+k_z\left(\frac{\partial^2 C}{\partial z^2}\right). \tag{2}$$

Both z and y are confined in the range $0<z<h$ and $0<y<L_y$ where h is the height of the planetary boundary layer (PBL) and L_y is a cross-wind distance faraway from the source, while the downwind distance $x>0$. The mathematical description of the dispersion problem (2) is completed by the following boundary conditions:

$$u\,C(x,y,z)=Q\,\delta(z)\,\delta(y)\,,\quad \text{at } x=0 \tag{3}$$

$$C(x,y,z)=0\,,\quad \text{at } x,y,z\to\infty \tag{4}$$

$$\frac{\partial C}{\partial y}=0\,,\quad \text{at } y=0,L_y \tag{5}$$

$$C(x,y,z)=R\,,\quad \text{at } y=0 \tag{6}$$

$$\frac{\partial C}{\partial z}=0\,,\quad \text{at } z=h \tag{7}$$

$$k_z\frac{\partial C}{\partial z}=-v_d\,C\,,\quad \text{at } z=0 \tag{8}$$

where v_d is the deposition velocity, Q is the emission rate and $R(x,z)$ is a variable.

The modeling of air pollution dispersion, including dry deposition, was first attempted by modifying the Gaussian plume equation (Chamberlain [2] and Overcamp [3]) and including operative algorithm, as in the surface depletion models (Horst [4, 5]). Ermak [6] found also an analytical solution but with diffusivity and wind as functions of down distance only and Berkowicz and Prahm [7] gave a numerical solution for the dependent time two dimensional equation including dry deposition. The solutions proposed by Smith [8] and Rao [9] also retained the framework of invariant wind speed and eddies with height (as the Gaussian approach). Tsuang [10] proposed a Gaussian model where the dispersion coefficients (the so-said "sigma") are functions of time and height.

Recent analytical solutions of the advection diffusion equation with dry deposition at the ground have utilized height-dependent wind speed and eddy diffusivities (Horst and Slinn [4], Koch [11], Chrysikopoulos et al. [12] and Tirbassi [13]). However, these solutions are restricted to the specific case in which the source is located at the ground level and/or with restrictions to the wind speed and eddy diffusivity vertical pro-

files. It is to be noted that the previous works, Moreira et al. [14, 15] assumed boundary conditions only of the second type (zero flux to the ground) and also Tirabassi et al. [16], but Tirabassi et al. [17] assumed boundary conditions of the the third kind (with deposition to the ground), which encompass the contaminant deposition speed and eddy, where eddy diffusivity profiles are functions in the z direction only.

The differential transform method is used in many fields and many mathematical physical problems such as a system of differential equations [18], a class of time dependent partial differential equations (PDEs) [19], wave, Laplace and heat equations [20], the fractional diffusion equations [21], two-dimensional transient heat flow [22], nonlinear partial differential equations [23], diffusion-convection equation [24], convection-dispersion problem [25], linear transport equation [26], two-dimension transient atmospheric pollutant dispersion [27], Helmholtz equation [28].

The aim of this work is to find the analytical solution developed for concentration of the pollutant released from an elevated source in an inversion layer by using the differential transform method (DTM) [29, 30] with different formulas of dispersion parameters (σ).

The paper is organized as follows. In section 2, we introduce the analytical solution using the differential transform method. In section 3, we apply both the standard method, power law, Briggs formula and other sigma to specific problems in analytical solution.

The validity of the present model is examined by comparing its results with the data for Cs^{137} which were performed around the Atomic Energy Authority (AEA) First Research Reactor in Egypt. The results are tabulated with the observed data and clarified in the conclusion.

2 Analytical solution

Applying DTM for (2) with respect to x, we get:

$$\frac{\partial U_i(x,y)}{\partial x} = k_y \frac{\partial^2 U_i(x,y)}{\partial y^2} + (i+1)(i+2)\, k_z U_{i+2}(x,y) \quad (9)$$

where the inverse of the differential transform is defined as:

$$C(x,y,y,z) = \sum_{i=0} z^i U_i(x,y)\,; \quad (10)$$

from boundary condition (8), we obtain:

$$U_1 = \left(\frac{-v_g}{k_z}\right) U_0\,; \quad (11)$$

from equations (9) and (11), we find that:

$$U_2 = \frac{1}{2k_z}\left(u\,\frac{\partial U_0(x,y)}{\partial x} - k_y \frac{\partial^2 U_0(x,y)}{\partial y^2}\right), \quad (12)$$

$$U_3 = \frac{-v_g}{6k_z}\left[u\,\frac{\partial}{\partial x}\left(\frac{U_0(x,y)}{k_z}\right) - \left(\frac{k_y}{k_z}\right)\frac{\partial^2 U_0(x,y)}{\partial y^2}\right]; \quad (13)$$

from boundary condition (7), we obtain:

$$\left(\frac{-2k_z v_g}{hk_y(2k_z - hv_g)}\right) U_0(x,y) + \left(\frac{u}{k_y}\right)\frac{\partial U_0(x,y)}{\partial x} - \\ -\left(\frac{2hk_z u v_g}{k_y(2k_z - hv_g)}\right) U_0(x,y)\,\frac{\partial}{\partial x}\left(\frac{1}{k_z}\right) = \frac{\partial^2 U_0(x,y)}{\partial y^2}\,. \quad (14)$$

By using separation of variables method for (14), we get:

$$\frac{d^2Y}{dy^2} + \lambda^2 Y = 0 \quad (15)$$

and

$$\frac{dX}{dx} - (A - B)X = 0 \quad (16)$$

where

$$A = \left(\frac{hk_z v_g \frac{\partial}{\partial x}\left(\frac{1}{k_z}\right)}{2k_z - hv_g}\right)$$

and

$$B = \frac{hk_z\lambda^2(2k_z - hv_g) - 2v_g k_z}{hu(2k_z - hv_g)}\,.$$

The solution of (14) becomes:

$$U_0(x,y) = c_1\, e^{\int (A-B)dx}\cos\lambda y \quad (17)$$

where $\lambda = n\pi/l_y$.

For practical application of solutions, we need to find the dispersion parameters σ_y, σ_z and the wind speed u. The dispersion parameters are an important function of downwind distance and stability. The empirical σ_y, σ_z curves suggested by Pasquill [31], Gifford [32] and Turner [33] have often been used and are based on the stability. There are different methods to find these parameters.

The meteorological conditions defining Pasquill turbulence types are

A- Extremely unstable conditions
B- Moderately unstable conditions
C- Slightly unstable conditions
D- Neutral conditions
E- Slightly stable conditions
F- Moderately stable conditions.

Here, we used four methods for estimating dispersion parameters:

1. Standard method: This method is based on a single atmospheric stability. Analytical expressions based on Pasquill-Gifford (P-G) curves used for the dispersion estimates have the forms [34]: .

$$\sigma_y = \frac{rx}{(1 + x/a)^p}, \quad (18)$$

$$\sigma_z = \frac{sx}{(1 + x/a)^q}, \quad (19)$$

where r, s, a, p and q are constants depending on the atmospheric stability. Table 1 shows the values of these constants for different stability classes [35].

Table 1: Meteorological data of the eight convective test runs [35]

Pasquil classes	A	B	C	D	E	F
σ_θ	25°	20°	15°	10°	5°	2.5°
$a(km)$	0.927	0.370	0.283	0.707	1.07	1.17
$s(m/km)$	102.0	96.2	72.2	47.5	33.5	22.0
q	-1.918	-0.101	0.102	0.465	0.624	0.70
$r(m/km)$	250	202	134	78.7	56.6	37.0
p	0.189	0.162	0.134	0.135	0.137	0.134

2. Power law of sigma: In this method σ_z and σ_y can be calculated from:

$$\sigma_y = cx^m \tag{20}$$

$$\sigma_z = dx^n . \tag{21}$$

The parameters c, d, m, n in Smith's (1968) [8] are estimated in table 2.

Table 2: Meteorological data of the eight convective test runs [36]

Pasquil classes	c	m	d	n
A-B	1.46	0.71	0.01	1.54
C	1.52	0.69	0.04	1.17
D	1.36	0.67	0.09	0.95
E-F	0.79	0.70	0.40	0.67

3. Briggs formulas: Formulas had been recommended by Briggs 1973 [37]; they should be used in place of the formulas in Table 3 to estimate σ_z and σ_y.

Table 3: Meteorological data of the eight convective test runs [35,37]

stability class	σ_y	σ_z
A-B	$0.32x\,(1+0.0004x)^{-\frac{1}{2}}$	$0.24x\,(1+0.001x)^{\frac{1}{2}}$
C	$0.22x\,(1+0.0004x)^{-\frac{1}{2}}$	$20x$
D	$0.16x\,(1+0.0004x)^{-\frac{1}{2}}$	$0.14x\,(1+0.0003x)^{-\frac{1}{2}}$
E-F	$0.11x\,(1+0.0004x)^{-\frac{1}{2}}$	$0.08x\,(1+0.00015x)^{-\frac{1}{2}}$

4. Hosker expression: Hosker 1973 [38] well-known analytical "best-fit" expression as:

$$\sigma_z = \left(\frac{\alpha x^\beta}{1+\gamma x^\delta}\right) F(z_0, x) \tag{22}$$

where z_0 is the roughness length, α, β, γ and δ are constants depending on the stability classes in Table 4 and $F(z_0, x)$ is defined as:

$$F(z_0, x) = \ln\left(mx^g\left[1+\left(lx^j\right)^{-1}\right]\right), \quad z_0 \geqslant 0.1m \tag{23}$$

where m, g, l, j are constants depend on the value of the roughness length, where our application z_0 (roughness length) = 0.5, so $l = 18.6$, $m = 5.16$, $j = 0.225$ and $g = 0.098$.

Table 4: The constant values of the roughness length, α, β, γ and δ [38]

Pasquil classes	α	β	γ	δ
A	0.112	1.06	5.0×10^{-4}	0.815
B	0.130	0.950	6.52×10^{-4}	0.750
C	0.112	0.920	9.05×10^{-4}	0.718
D	0.098	0.889	1.35×10^{-3}	0.688
E	0.0609	0.895	1.96×10^{-3}	0.684
F	0.0638	0.783	1.36×10^{-3}	0.672

On the other hand, Briggs 1973 [37] proposed a series of algebraic interpolation formulae based on a wide variety of data sources containing surface and elevated sources with a range of initial buoyancies:

$$\sigma_y = b_1(1+b_2x)^{b_3} . \tag{24}$$

The coefficient values b_1, b_2 and b_3 were derived for both rural and urban terrain and are given in Table 5 [37].

Table 5: The coefficient values b_1, b_2 and b_3 for equation (24) [37]

PG stability	b_1	b_2	b_3
A	0.20	0	–
B	0.12	0	–
C	0.08	0.0002	-0.5
D	0.06	0.0015	-0.5
E	0.03	0.0003	-1
F	0.016	0.0003	-1

3 Results and discussion

Meteorological data provided by Inshas meteorological tower for four months at a smooth flat site (Inshas area, Egypt) for the year (2006) are given in Table 6, [39]. Air samples were collected from 98 m to 186 m around the first and second research reactor in AEA, Egypt. The study area is flat, dominated by sand soil with poor vegetation cover. The study area was divided into 16 sectors (with 22.5° width for each sector), beginning from the north direction. Aerosols were collected at a height of 0.7 m above the ground of 10.3 cm diameter filter paper with a desired collection efficiency (3.4%)

Table 6: Meteorological Data of the nine Convective test runs at Inshas Site

No.	Stability	Down distance x (m)	Mixing height (m)	Emission rate Q (Bq)	Wind speed (m/s)	Initial wind velocity u_0(m/s)
1	A	98	600.85	0.555429	4	3.95
2	A	100	801.13	0.567	4	3.7
3	B	106	973.0	0.023143	6	5.1
4	C	106	888.0	0.254577	4	3.95
5	A	135	921.0	0.266143	4	3.1
6	D	136	443.0	0.277714	4	3.95
7	E	154	1271.0	0.543857	4	3.95
8	C	165	1842.0	0.563529	4	3.1
9	A	186	1642	0.558321	4	3.95

Table 7: Observed and calculated concentrations (Bq/m^3) for nine experiments

Run no.	Observed Con. [39]	Calculated concentrations			
		Standard Model	Power law of sigma	Briggs formulas	Hosker expression
1	0.002	0.0140799	0.0189143	0.013563	0.01379
2	0.004	0.0153392	0.011873	0.01475	0.014014
3	0.005	0.00448	0.00507518	0.004391	0.04422
4	0.007	0.0062904	0.013799	0.00624019	0.00625
5	0.009	0.00859466	0.00870	0.0081565	0.0081117
6	0.007	0.0070497	0.01596	0.0068969	0.00674
7	0.007	0.0137824	0.015015	0.019399	0.013155
8	0.019	0.0177893	0.019194	0.0171672	0.017135
9	0.006	0.0141444	0.01312	0.0132115	0.01311

using a high volume air sampler with 220 V / 50 Hz bias. The air sampler had an air flow rate of approximately 0.7 m^3/min (25 ft^3/min). Sample collective time was 30 min with an air volume of 21.2 m^3 (750 ft^3). This air volume was corrected to standard conditions (25 C^o and 1013 mb) [39].

Table 7 indicates comparison between experimental data of the nine convective test runs at Inshas site and our calculation of concentration by Briggs formula, power law variation, standard method and Hosker's expression, which shows that the power law formula for the dispersion coefficients achieves the best agreement with the experimental results.

3.1 Statistical evaluation

Statistical analysis of the predictions and observations is central to the model performance evaluation. The predicted and the corresponding observed concentrations are treated as pairs in this evaluation.

The statistical index FB indicates weather the predicted quantities underestimate or overestimate the observed ones. The statistical index NMSE represents the quadratic error of the predicted quantities in relation to the observed ones. Best results are indicated by values nearest zero in NMSE, FB, nearest 1 in MG, VG and FAC2 and are factor of two if are greater than 1 and less than 2. The statistical measures chosen to compare performances of the models described here [40]:
(i) Fractional bias FB is defined as:

$$\mathrm{FB} = \frac{\bar{C}_o - \bar{C}_p}{0.5(\bar{C}_o + \bar{C}_p)}$$

where the subscripts o and p refer to the observed and predicted values, respectively, and the overbars indicate mean values. A good model should have FB value close to zero.
(ii) Normalized mean square error (NMSE) is defined as:

$$\mathrm{NMSE} = \frac{\overline{(C_o - C_p)^2}}{\bar{C}_o \bar{C}_p} .$$

This provides information on the overall deviations between predicted and observed concentrations. It is a dimensionless statistic and its value should be as small as possible for a good model.
(iii) The geometric mean bias is defined as:

$$MG = \exp\left(\overline{\ln C_o} - \overline{\ln C_p}\right) .$$

(iv) The geometric variance is defined as:

$$VG = \exp\left(\overline{(\ln C_o - \ln C_p)^2}\right) .$$

Table 8: Comparison between the Standard method, Power law of sigma, Briggs formulas and Hosker expression in terms of FB, FAC2, NMSE, MG and VG

	Standard method	Power law of sigma	Briggs formulas	Hosker expression
FB	-0.42543	-0.65368	-0.445	-0.69646
FAC2	1.540371	1.971068	1.572352	2.068563
NMSE	0.189565	0.478407	0.208342	0.551991
MG	0.605381	0.46362	0.601944	0.490099
VG	1.286468	1.805587	1.293883	1.662927

(v) Fraction within a factor of two (FAC2) is given by:

$$0.5 \leqslant (C_p/C_o) \leqslant 2\,.$$

Statistical evaluation of the models results are given in Table 8, which compares the Standard method, Power law of sigma, Briggs formulas and Hosker expression in terms of FB, FAC2, NMSE, MG and VG.

4 Conclusion

In the present study, an analytical treatment for the dispersion of air pollutant released from point source is formulated. A mathematical solution has been obtained for the steady-state form of the three-dimensional advection-diffusion equation using the Differential Transform Method. Different realistic formulae for the dispersion coefficients as a function of downwind distance have been adopted (namely: Briggs formula, power law variation, standard method and Hosker's expression). In order to validate and verify our model, and for the sake of comparison, we apply our obtained mathematical formulae on the experimental data performed for the release from the first Research Reactor in Egypt. The comparison shows that the power law formula for the dispersion coefficients achieves the best agreement with the experimental results. Finally, the good agreement between the power law variation of the dispersion parameter and the experiential data gives us confidence to extend this work for the case of different sources types, namely, line, area and volume sources. In addition, it is also our intention to perform the mathematical analysis of this method for the case of high penetrated inversion layer (i.e. different stability conditions that permits the pollutant penetration and diffusion through the mixing height).

Submitted on January 25, 2016 / Accepted on February 11, 2016

References

1. Blackadar A. Turbulence and Diffusion in the Atmosphere: Lectures in Environmental Sciences. Springer-Verlag, 1997.
2. Chamberlain A. Aspects of Travel and Deposition of Aerosol and Vapour Clouds. UKAEA Rep. AERE-HP/R-1261, Harwell, Berkshire, United Kingdom, 1953.
3. Overcamp T. A general Gaussian diffusion deposition model for elevated point sources. *J. Appl. Meteor.*, 1976, v. 15, 1167–1171.
4. Horst T., Slinn W. G. A surface depletion model for deposition from a Gaussian plume. *Atmos. Environ.*, 1984, v. 18, 1339–1346.
5. Horst T. The modification of plume models to account for dry deposition. *Bound. Layer Meteor.*, 1984, v. 30, 413–430.
6. Ermak D. An analytical model for air polutant transport and deposition from a point source. *Atmos. Environ.*, 1976, v. 11, 231–237.
7. Berkowicz R. and Prahm L. Pseudospectral simulation of dry deposition from a point source. *Atmos. Environ.*, 1977, v. 12, 379–387.
8. Smith F. The problem of deposition in atmospheric diffusion of particulate matter. *J. Atmos. Sci.*, 1962, v. 19, 429–434.
9. Rao K. Analytical Solutions of a Gradient-transfer Model for Plume Deposition and Sedimentation. NOAA Tech. Memo. ERL ARL-109, Air Resources Laboratories, 1981.
10. Tsuang B. A Gaussian plume trajectory model to quantify the source/receptor relationship of primary pollutants and secondary aerosols: Part I. Theory. *Atmos. Environ.*, 2003, v. 37, 3981–3991.
11. Koch W. A solution of two-dimensional atmospheric diffusion equation with height-dependent diffusion coefficient including ground level absorption. *Atmos. Environ.*, 1989, v. 23, 1729–1732.
12. Chrysikopoulos C., Hildemann L. and Roberts P. A three-dimensional steady-state atmospheric dispersion deposition model for emissions from a ground-level area source. *Atmos. Environ.*, 1992, v. 26A, 747–757.
13. Tirabassi T. Operational advanced air pollution modeling. *PAGEOPH*, 2003, v. 160, 5–16.
14. Moreira D., Rizza U., Vilhena M. and Goulart A. Semi-analytical model for pollution dispersion in the planetary boundary layer. *Atmos. Environ.*, 2005, v. 39, 2689–2697.
15. Moreira D., Vilhena M., Tirabassi T., Costa C. and Bodmann B. Simulation of pollutant dispersion in atmosphere by the Laplace transform: the ADMM approach. *Water Air Soil Pollut.*, 2006, v. 177, 411–439.
16. Tirabassi T., Buske D., Moreira D. and Vilhena M. A two-dimensional solution of the diffusion equation with dry deposition to the ground. *J. Appl. Meteorol. Climatol.*, 2008, v. 47, 2096–2104.
17. Tirabassi T., Moreira D., Vilhena M. and Goulart A. A multi-layer model for pollutants dispersion with dry deposition to the ground. *Atmos. Environ.*, 2010, v. 44, 1859–1865.
18. Thongmoon M. and Pusjuso S. The numerical solutions of differential transform method and the Laplace transform method for a system of differential equations. *Nonlinear Anal. Hybrid Sys.*, 2010, v. 4, 425–431.
19. Kong W. and Xionghua Wu. The Laplace transform and polynomial Trefftz method for a class of time dependent PDEs. *Appl. Math. Model.*, 2009, v. 33, 2226–2233.
20. Eltayeb H. and Kiliçman A. A note on solutions of wave, Laplace's and heat equations with convolution terms by using a double Laplace transform. *Appl. Math. Let.*, 2008, v. 21, 1324–1329.
21. Zamani K. One-dimensional, mass concervative, spatially dependent transport equation: new analytical solution. 12th Pan-American Congress of Applied Mechanics, 2012, 1–6.
22. Mahajerin E. and Burgess G. A Laplace transform-based fundamental collocation method for two-dimensional transient heat flow. *Appl. Therm. Eng.*, 2003, v. 23, 101–111.

23. Abassy T., El-Tawil M. and El-Zoheiry H. Exact solutions of some nonlinear partial differential equations using the variational iteration method linked with Laplace transforms and the Padé technique. *Comput. Math. Appl.*, 2007, v. 54, 940–954.
24. Davis G. A Laplace transform technique for the analytical solution of a diffusion-convection equation over a finite domain. *Appl. Math. Model.*, 1985, v. 9, 69–71.
25. Ren L. and Zhang R. Hybrid Laplace transform finite element method for solving the convection-dispersion problem. *Advan. Wat. Res.*, 1999, v. 23, 229–237.
26. Cardona A. and Vilhena M. A solution of the linear transport equation using Walsh function and Laplace transform. *Annal. Nucl. Energ.*, 1994, v. 21, 495–505.
27. Cassol M., Wortmann S. and Rizza U. Analytic modeling of two-dimensional transient atmospheric pollutant dispersion by double GITT and Laplace Transform techniques. *Environ. Model. Software*, 2009, v. 24, 144–151.
28. Souchet R. Laplace transform method in boundary value problems for the Helmholtz equation. *Mech. Res. Commun.*, 1980, v. 7, 159–164.
29. Raslan K. R., Biswas A. and Abou Sheer Z. F. *Int. J. Phys. Sci.*, 2012, v. 7 (9), 1412–1419.
30. Sohail M. and Mohyud-Din S. T. *Int. J. Modern Eng. Sci.*, 2012, v. 32 (1), 14–22.
31. Pasquill F. *Meteorol. Mag.*, 1961, v. 90, 33–49.
32. Gifford F. A. *Nucl. Safety*, 1961, v. 2 (4), 47–57.
33. Turner D. B. *J. Air Poll. Control Assoc.*, 1979, v. 29, 502–519.
34. Green A. E. S., Singhal R. P. and Venkateswar R. Analytic extensions of the Gaussian plume model. *J. Air Poll. Control Assoc.*, 1980, v. 30 (7), 773–776.
35. Essa K., Mubarak F. and Abu Khadra S. Comparison of some sigma schemes for estimation of air pollutant dispersion in moderate and low winds. *Atmos. Sci. Let.*, 2005, v. 6, 90–96.
36. Panofsky H. A., Dutton J. A. Atmospherical Turbulence Models and Methods for Engineering Applications. John Wiley & Sons, New York, 1984.
37. Briggs G. A. Diffusion Estimation for Small Emissions. ATDL Contribution File No. 79, 1973.
38. Hosker R. Estimates of Dry Deposition and Plume Depletion over Forests and Grasslands in Physical Behaviour or Radioactive Contaminats in the Atmosphere. Symposium Proceedings, Vienna International Atomic Energy Agency, Vienna, 1973, 291–308.
39. Halfa I. K. I. Analysis of the Experimental Data of Air Pollution Using Atmospheric Dispersion Modeling and Rough Set, Doctoral Thesis, Faculty of Science, Tanta University, 2008.
40. Hanna S., Hansen O., and Dharmavaram S. FLACS CFD Air Quality Model Performance Evaluation with Kit Fox, MUST, Prairie Grass, and EMU Observations. *Atmos. Environ.*, 2004, v. 38, 4675–4687.

On the Applicability of Bell's Inequality

Pierre A. Millette
PierreAMillette@alumni.uottawa.ca, Ottawa, Canada

We investigate the applicability of Bell's inequality based on the assumptions used in its derivation. We find that it applies to a specific class of hidden variable theories referred to as Bell theories, but not necessarily to other hidden variable dynamic theories. We consider examples of quantum dynamical processes that cannot be represented by the initial representation defined in Bell's derivation. We highlight two hidden assumptions identified by Jaynes [11] that limit the applicability of Bell's inequality, as derived, to Bell hidden variable theories and that show that there are no superluminal physical influences, only logical inferences.

1 Introduction

Bell's inequality [1–3] sets constraints for the existence of local hidden variable theories in quantum mechanics. Bohr, of the Copenhagen probabilistic school, and Einstein, of the objective reality school, who both contributed to the foundation of quantum mechanics, did not agree on its interpretation – their views and correspondence on the topic are well documented in many books [4–7].

In 1935, Einstein, Podolsky and Rosen published a paper [8] that aimed to show that quantum mechanics was not a complete description of physical reality. Bohr provided a response to the challenge [9], but the EPR paper remained an argument for hidden variables in quantum mechanics. In 1964, Bell [1] published an inequality that imposed constraints for local hidden variable theories to be valid in quantum mechanics. The experiments performed by Aspect *et al* [10] with entangled photons confirmed that Bell's inequality was violated within experimental errors, taken to mean that local hidden variable theories are not valid in quantum mechanics. Only non-local hidden variable theories are possible, based on these results.

In this paper, we investigate the applicability of Bell's inequality, based on the assumptions used in its derivation.

2 Bell's inequality

Bell's derivation [1] considers a pair of spin one-half particles of spin σ_1 and σ_2 respectively, formed in the singlet state, and moving freely in opposite directions. Then $\sigma_1 \cdot \mathbf{a}$ is the measurement of the component of σ_1 along some vector $\mathbf{a}$, and similarly for $\sigma_2 \cdot \mathbf{b}$ along some vector $\mathbf{b}$. Bell then considers the possibility of a more complete description using hidden variable parameters λ.

He writes down the following equation for the expectation value of the product of the two components $\sigma_1 \cdot \mathbf{a}$ and $\sigma_2 \cdot \mathbf{b}$ with parameters λ:

$$P(\mathbf{a}, \mathbf{b}) = \int d\lambda \, \rho(\lambda) \, A(\mathbf{a}, \lambda) \, B(\mathbf{b}, \lambda) \tag{1}$$

where

$$A(\mathbf{a}, \lambda) = \pm 1 \text{ and } B(\mathbf{b}, \lambda) = \pm 1 \tag{2}$$

and $\rho(\lambda)$ is the probability distribution of parameter λ. This should equal the quantum mechanical expectation value

$$<\sigma_1 \cdot \mathbf{a} \; \sigma_2 \cdot \mathbf{b}> = -\mathbf{a} \cdot \mathbf{b} \, . \tag{3}$$

Bell says that it does not matter whether λ is "a single variable or a set, or even a set of functions, and whether the variables are discrete or continuous" [1]. He uses a single continuous parameter described by a probability distribution. In a later paragraph, he states that (1) represents all kinds of possibilities, such as any number of hidden variables, two sets of hidden variables dependent on A and B, or even as initial values of the variables λ at a given time if one wants to assign "dynamical significance and laws of motion" [1] to it. However, it is doubtful that the probability distribution $\rho(\lambda)$ can be used to represent all possible theories of hidden variables.

Indeed, the basic limitation of (1) with its use of a probability distribution $\rho(\lambda)$ is that it imposes a quantum mechanical calculation representation on the analysis. Other quantum level dynamic theories, which we will refer to as hidden variable dynamic theories, could obey totally different dynamic principles, in which case, (1) would not be applicable. Equation (1) is only applicable to a specific class of hidden variable theories that can be represented by that equation, which Jaynes [11] refers to as Bell theories. In the following sections, we consider examples of quantum dynamical processes that cannot be represented by (1) or by the probability distribution $\rho(\lambda)$ used in (1).

3 Measurement limitations and inherent limitations

It is important to note that Bohr's responses to Einstein's *gedanken* experiments were based on measurements arguments, which acted as a barrier to any further analysis beyond that consideration. As pointed out by Jaynes [12], Einstein and Bohr "were both right in the essentials, but just thinking on different levels. Einstein's thinking [was] always on the ontological level traditional in physics; trying to describe the realities of Nature. Bohr's thinking [was] always on the epistemological level, describing not reality but only our information about reality".

As discussed in [13], the Heisenberg Uncertainty Principle arises because x and p form a Fourier transform pair of variables at the quantum level due to the momentum p of a quantum particle being proportional to the de Broglie wave number k of the particle. It is a characteristic of quantum mechanics that conjugate variables are Fourier transform pairs of variables.

It is thus important to differentiate between the measurement limitations that arise from the properties of Fourier transform pairs, and any inherent limitations that may or may not exist at the quantum level for those same variables, independently of the measurement process. Conjugate variable measurement limitations affect how we perceive quantum level events as those can only be perceived by instrumented measurements at that level. However, as shown in [13], conjugate variable measurement limitations affect *only* our perception of the quantum environment, and are *not* inherent limitations of the quantum level.

The Nyquist-Shannon Sampling Theorem of Fourier transform theory allows access to the range of values of variables below the Heisenberg Uncertainty Principle limit under sampling measurement conditions, as demonstrated by the Brillouin zones formulation of solid state physics [13] [14, see p. 21] [15, see p. 100]. Physically this result can be understood from the sampling measurement operation building up the momentum information during the sampling process, up to the Nyquist limit. This shows that there are local hidden variables at the quantum level, independently of the measurement process. The dynamical process in this case is masked by the properties of the Fourier transform.

4 Wave-particle duality in *STCED*

The Elastodynamics of the Spacetime Continuum (*STCED*) [16] has similarities to Bohmian mechanics in that the solutions of the *STCED* wave equations are similar to Louis de Broglie's "double solution" [17, 18]. Bohmian mechanics also known as de Broglie-Bohm theory [19–21] is a theory of quantum physics developed by David Bohm in 1952 [22], based on Louis de Broglie's original work on the *pilot wave*, that provides a causal interpretation of quantum mechanics. It is empirically equivalent to orthodox quantum mechanics, but is free of the conceptual difficulties and the metaphysical aspects that plague the interpretation of quantum theory.

Interestingly, Bell was aware of and a proponent of Bohmian mechanics when he derived his inequality [23]:

> "Bohm showed explicitly how parameters could indeed be introduced, into nonrelativistic wave mechanics, with the help of which the indeterministic description could be transformed into a deterministic one.More importantly, in my opinion, the subjectivity of the orthodox version, the necessary reference to the 'observer,' could be eliminated... I will try to present the essential idea... so compactly, so lucidly, that even some of those who know they will dislike it may go on reading, rather than set the matter aside for another day."

In Bohmian mechanics, a system of particles is described by a combination of the wavefunction from Schrodinger's equation and a guiding equation that specifies the location of the particles. "Thus, in Bohmian mechanics the configuration of a system of particles evolves via a deterministic motion choreographed by the wave function" [21] such as in the two-slit experiment. We will see a similar behavior in the *STCED* wave equations below. Bohmian mechanics is equivalent to a non-local hidden variables theory.

In the Elastodynamics of the Spacetime Continuum, as discussed in [24], energy propagates in the spacetime continuum by longitudinal (*dilatation*) and transverse (*distortion*) wave displacements. This provides a natural explanation for wave-particle duality, with the transverse mode corresponding to the wave aspects of the deformations and the longitudinal mode corresponding to the particle aspects of the deformations.

The displacement u^ν of a deformation from its undeformed state can be decomposed into a longitudinal component $u^\nu_\parallel$ and a transverse component $u^\nu_\perp$. The volume dilatation ε is given by the relation $\varepsilon = u_\parallel{}^\mu{}_{;\mu}$ [16]. The wave equation for $u^\nu_\parallel$ describes the propagation of longitudinal displacements, while the wave equation for $u^\nu_\perp$ describes the propagation of transverse displacements in the spacetime continuum. The u^ν displacement wave equations can be expressed as a longitudinal wave equation for the dilatation ε and a transverse wave equation for the rotation tensor $\omega^{\mu\nu}$ [16].

Particles propagate in the spacetime continuum as longitudinal wave displacements. Mass is proportional to the volume dilatation ε of the longitudinal mode of the deformation [16, see (32)]. This longitudinal mode displacement satisfies a wave equation for ε, different from the transverse mode displacement wave equation for $\omega^{\mu\nu}$. This longitudinal dilatation wave equation for ε is given by [16, see (204)]

$$\nabla^2\varepsilon = -\frac{\bar{k}_0}{2\bar{\mu}_0 + \bar{\lambda}_0}\, u^\nu_\perp \varepsilon_{;\nu} \tag{4}$$

where $\bar{\mu}_0$ and $\bar{\lambda}_0$ are the Lamé constants and $\bar{k}_0$ the elastic volume force constant of the spacetime continuum. It is important to note that the inhomogeneous term on the R.H.S. includes a dot product coupling between the transverse displacement and the volume dilatation for the solution of the longitudinal dilatation wave equation for ε.

The transverse distortion wave equation for $\omega^{\mu\nu}$ [16, see (210)]

$$\nabla^2\omega^{\mu\nu} + \frac{\bar{k}_0}{\bar{\mu}_0}\,\varepsilon(x^\mu)\;\omega^{\mu\nu} = \frac{1}{2}\frac{\bar{k}_0}{\bar{\mu}_0}\,(\varepsilon^{;\mu}u^\nu_\perp - \varepsilon^{;\nu}u^\mu_\perp) \tag{5}$$

also includes a R.H.S. coupling, in this case a cross product, between the transverse displacement and the volume dilatation for the solution of the transverse distortion wave equa-

tion for $\omega^{\mu\nu}$. The transverse distortion wave $\omega^{\mu\nu}$ corresponds to a multi-component wavefunction Ψ.

A deformation propagating in the spacetime continuum consists of a combination of longitudinal and transverse waves. The coupling between $\varepsilon^{,\mu}$ and $u^{\nu}_{\perp}$ on the R.H.S. of both wave equations explains the behavior of electrons in the double slit interference experiment. It shows that even though the transverse wave is the source of the interference pattern in double slit experiments, the longitudinal dilatation wave, which behaves as a particle, follows the interference pattern dictated by the transverse distortion wave as observed experimentally. The longitudinal dilatation wave behaves as a particle and goes through one of the slits, even as it follows the interference pattern dictated by the transverse distortion wave, as observed experimentally [25, see in particular Figure 4] and as seen in the coupling between $\varepsilon^{,\mu}$ and $u^{\nu}_{\perp}$ in (4) and (5) above. This behavior is the same as that in Bohmian mechanics seen above. These results are in agreement with the results of the Jánossy-Naray, Clauser, and Dagenais and Mandel experiments on the self-interference of photons and the neutron interferometry experiments performed by Bonse and Rauch [26, see pp. 73-81].

As mentioned previously, the solutions of the *STCED* wave equations are similar to Louis de Broglie's "double solution". The longitudinal wave is similar to the de Broglie "singularity-wave function" [17]. In *STCED* however, the particle is not a singularity of the wave, but is instead characterized by its mass which arises from the volume dilatation ε propagating as part of the longitudinal wave. There is no need for the collapse of the wavefunction Ψ, as the particle resides in the longitudinal wave, not the transverse one. A measurement of a particle's position is a measurement of the longitudinal wave, not the transverse wave.

In addition, $|\Psi|^2$ represents the physical energy density of the transverse (*distortion*) wave. It corresponds to the transverse field energy of the deformation. It is not the same as the particle, which corresponds to the longitudinal (*dilatation*) wave displacement and is localized within the deformation via the massive volume dilatation. However, $|\Psi|^2$ can be normalized with the system energy and converted into a probability density, thus allowing the use of the existing probabilistic formulation of quantum theory.

The dynamical process, although it has some similarities to Bohmian mechanics, is also different from it as it is centered on longitudinal (particle) and transverse (wavefunction) wave equations derived from the properties of the spacetime continuum of general relativity. It is thus deterministic and causal as is general relativity.

5 Physical influence versus logical inference

We have considered two examples of quantum dynamical processes where the starting equation (1) and the probability distribution $\rho(\lambda)$ used in (1) do not apply to the situation. We now examine in greater details the probabilistic formulation of Bell's inequality derivation of section 2 to better understand its limitations.

Physicist E. T. Jaynes was one of the proponents of the usage of probability theory as an extension of deductive logic. His textbook "Probability Theory: The Logic of Science" [27] published posthumously is an invaluable resource for scientists looking to understand the scientific use of probability theory as opposed to the conventional mathematical measure theory. As he states in [11],

> "Many circumstances seem mysterious or paradoxical to one who thinks that probabilities are real physical properties existing in Nature. But when we adopt the "Bayesian Inference" viewpoint of Harold Jeffreys [28,29], paradoxes often become simple platitudes and we have a more powerful tool for useful calculations."

Jaynes clarifies this approach to probability theory and contrasts it to frequencies as follows [11]:

> "In our system, a probability is a theoretical construct, on the epistemological level, which we assign in order to represent a state of knowledge, or that we calculate from other probabilities according to the rules of probability theory. A frequency is a property of the real world, on the ontological level, that we measure or estimate."

The probability distributions used for inference do not describe a property of the world, only a certain state of information about the world, which provides us with the means to use prior information for analysis as powerfully demonstrated in numerous applications in [11, 12, 27].

The Einstein–Podolsky–Rosen (EPR) paradox and Bell inequality in quantum theory is one of the examples examined by Jaynes in [11]. In quantum mechanics, the belief that probabilities are real physical properties leads to quandaries such as the EPR paradox which lead some to conclude that there is no real world and that physical influences travel faster than the speed of light, or worse ("a spooky kind of action at a distance" as Einstein called it). As Jaynes points out, it is important to note that the EPR article did not question the existence of the correlations, which were expected, but rather the need for a physical causation instead of what he calls "instantaneous psychokinesis", based on experimenter decisions, to control distant events.

Jaynes' analysis of the derivation of Bell's inequality uses the following notation for conditional probabilities which corresponds to Bell's notation as follows:

$$P(AB\,|\,ab) = P(\mathbf{a},\mathbf{b}) \tag{6}$$

$$P(A\,|\,a\lambda) = A(\mathbf{a},\lambda)\,, \tag{7}$$

such that Bell's equation (1) above becomes

$$P(AB\,|\,ab) = \int d\lambda\,\rho(\lambda)\,P(A\,|\,a\lambda)\,P(B\,|\,b\lambda)\,. \tag{8}$$

However, as Jaynes notes, the fundamentally correct relation for $P(AB\,|\,ab)$ according to probability theory should be

$$P(AB\,|\,ab) = \int d\lambda\, P(AB\,|\,ab\lambda)\, P(\lambda\,|\,ab)\,. \qquad (9)$$

Assuming that knowledge of the experimenters' choices gives no information about λ, then one can write

$$P(\lambda\,|\,ab) = \rho(\lambda)\,. \qquad (10)$$

The fundamentally correct factorization of the other probability factor of (9), $P(AB\,|\,ab\lambda)$, is given by [11]

$$P(AB\,|\,ab\lambda) = P(A\,|\,ab\lambda)\, P(B\,|\,Aab\lambda)\,. \qquad (11)$$

However, as Jaynes notes, one could argue as Bell did that EPR demands that A should not influence events at B for space-like intervals. This requirement then leads to the factorization used by Bell to represent the EPR problem

$$P(AB\,|\,ab\lambda) = P(A\,|\,a\lambda)\, P(B\,|\,b\lambda)\,. \qquad (12)$$

Nonetheless, the factorization (12) disagrees with the formalism of quantum mechanics in that the result of the measurement at A must be known before the correlation affects the measurement at B, *i.e.* $P(B\,|\,Aab)$. Hence it is not surprising that Bell's inequality is not satisfied in systems that obey quantum mechanics.

Two additional hidden assumptions are identified by Jaynes in Bell's derivation, in addition to those mentioned above:

1. Bell assumes that a conditional probability $P(X\,|\,Y)$ represents a physical causal influence of Y on X. However, consistency requires that conditional probabilities express logical inferences not physical influences.
2. The class of Bell hidden variable theories mentioned in section 2 does not include all local hidden variable theories. As mentioned in that section, hidden variable theories don't need to satisfy the form of (1) (or alternatively (8)), to reproduce quantum mechanical results, as evidenced in Bohmian mechanics.

Bell's inequality thus applies to the class of hidden variable theories that satisfy his relation (1), *i.e.* Bell hidden variable theories, but not necessarily to other hidden variable dynamic theories.

The superluminal communication implication stems from the first hidden assumption above which shows that what is thought to travel faster than the speed of light is actually a logical inference, not a physical causal influence. As summarized by Jaynes [11],

> "The measurement at A at time t does not change the real physical situation at B; but it changes our state of knowledge about that situation, and therefore it changes the predictions we are able to make about B at some time t'. Since this is a matter of logic rather than physical causation, there is no action at a distance and no difficulty with relativity."

There is simply no superluminal communication, as required by special relativity. Assuming otherwise would be similar to Pauli assuming that the established law of conservation of energy mysteriously fails in weak interactions instead of successfully postulating a new particle (the neutrino).

6 Discussion and conclusion

In this paper, we have investigated the applicability of Bell's inequality, based on the assumptions used in its derivation. We have considered two examples of hidden variable dynamic theories that do not satisfy Bell's initial equation (1) used to derive his inequality, and consequently for which Bell's inequality is not applicable: one based on the Nyquist-Shannon Sampling Theorem of Fourier transform theory and the other based on the wave-particle solutions of the *STCED* wave equations which are similar to Louis de Broglie's "double solution". We highlight two hidden assumptions identified by Jaynes [11] that limit the applicability of Bell's inequality, as derived, to Bell hidden variable theories and that show that there are no superluminal physical influences, only logical inferences.

We close with a quote from Jaynes [27, see p. 328] that captures well the difficulty we are facing:

> "What is done in quantum theory today... when no cause is apparent one simply postulates that no cause exists – ergo, the laws of physics are indeterministic and can be expressed only in probability form."

Thus we encounter paradoxes such as seemingly superluminal physical influences that contradict special relativity, and "spooky action at a distance" is considered as an explanation rather than working to understand the physical root cause of the problem. This paper shows that, in this case, the root cause is due to improper assumptions, specifically the first hidden assumption identified by Jaynes highlighted in section 5 above, that is assuming that a conditional probability represents a physical influence instead of the physically correct logical inference. In summary,

> "He who confuses reality with his knowledge of reality generates needless artificial mysteries." [11]

Submitted on February 19, 2016 / Accepted on February 22, 2016

References

1. Bell J. S. On the Einstein–Podolsky–Rosen Paradox. *Physics*, 1964, v. 1, 195–200. Reprinted in Bell J. S. Speakable and Unspeakable in Quantum Mechanics. Cambridge University Press, Cambridge, 1987, pp. 14–21.
2. Goldstein S., Norsen, T., Tausk D. V., Zanghi, N. Bell's Theorem. *Scholarpedia*, 2011, v. 6 (10), 8378.
3. Gouesbet G. Hidden Worlds in Quantum Physics. Dover Publications, New York, 2013, p. 280–303.
4. Mehra J. Einstein, Physics and Reality. World Scientific Publishing, Singapore, 1999.

5. Whitaker A. Einstein, Bohr and the Quantum Dilemma; From Quantum Theory to Quantum Information, 2nd ed. Cambridge University Press, Cambridge, 2006.
6. Home D., Whitaker A. Einstein's Struggles with Quantum Theory; A Reappraisal. Springer, New York, 2007.
7. Stone A. D. Einstein and the Quantum; The Quest of the Valiant Swabian. Princeton University Press, Princeton, 2013.
8. Einstein A, Podolsky B. and Rosen N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.*, 1935, v. 47, 777–780.
9. Bohr N. Can Quantum Mechanical Description of Reality Be Considered Complete? *Phys. Rev.*, 1935, v. 48, 696.
10. Aspect A, Dalibard J. and Roger G. Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.*, 1982, v. 49, 1804–1807.
11. Jaynes E. T. Clearing Up Mysteries – The Original Goal. In Skilling J., ed. Proceedings Volume, Maximum Entropy and Bayesian Methods. Kluwer Academic Publishers, Dordrecht, 1989, pp. 1–27.
12. Jaynes E. T. Probability in Quantum Theory. In Zurek W. H., ed. Complexity, Entropy and the Physics of Information. Addison Wesley Publishing, Reading, MA, 1990.
13. Millette P. A. The Heisenberg Uncertainty Principle and the Nyquist-Shannon Sampling Theorem. *Progress in Physics*, 2013, v. 9 (3), 9–14. arXiv: quant-ph/1108.3135.
14. Ziman J. M. Principles of the Theory of Solids, 2nd ed. Cambridge University Press, Cambridge, 1979.
15. Chaikin P. M. and Lubensky T. C. Principles of Condensed Matter Physics. Cambridge University Press, Camvridge, 1995.
16. Millette P. A. Elastodynamics of the Spacetime Continuum. *The Abraham Zelmanov Journal*, 2012, v. 5, 221–277.
17. de Broglie L. Non-Linear Wave Mechanics. Elsevier Publishing, Amsterdam, 1960.
18. de Broglie L. Les incertitudes d'Heisenberg et l'interprétation probabiliste de la Mécanique Ondulatoire. Gauthier-Villars, Paris, 1982. Available in English: Heisenberg's Uncertainties and the Probabilistic Interpretation of Wave Mechanics. Kluwer Academic Publishers, Dordrecht, 1990.
19. Holland P. R. The Quantum Theory of Motion; An Account of the de Broglie-Bohm Causal Interpretation of Quantum Mechanics. Cambridge University Press, Cambridge, 1993.
20. Dürr D., Teufel, S. Bohmian Mechanics; The Physics and Mathematics of Quantum Theory. Springer-Verlag, Berlin, 2009.
21. Goldstein S. Bohmian Mechanics. *The Stanford Encyclopedia of Philosophy*, 2013, http://plato.stanford.edu/archives/spr2013/entries/qm-bohm/.
22. Bohm D. A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables I, II. *Physical Review*, 1952, v. 85, 166-–179, 180–193.
23. Bell J. S. On the Impossible Pilot Wave. *Foundations of Physics*, 1982, v. 12, 989–999. Reprinted in Bell J. S. Speakable and Unspeakable in Quantum Mechanics. Cambridge University Press, Cambridge, 1987, pp. 159–168.
24. Millette P. A. Wave-Particle Duality in the Elastodynamics of the Spacetime Continuum (STCED). *Progress in Physics*, 2014, vol. 10 (4), 255–258.
25. Hasselbach F. Recent Contributions of Electron Interferometry to Wave-Particle Duality. In Selleri F., ed. Wave-Particle Duality. Plenum, New York, 1992, 109–125.
26. Selleri F. Quantum Paradoxes and Physical Reality. Kluwer Academic Publishers, Dordrecht, 1990.
27. Jaynes E. T., Bretthorst G. L., ed. Probability Theory; The Logic of Science. Cambridge University Press, Cambridge, 2003.
28. Jeffreys H. Scientific Inference. Cambridge University Press, Cambridge, 1931.
29. Jeffreys H. Theory of Probability. Oxford University Press, Oxford, 1939.

Coincident Down-chirps in GW150914 Betray the Absence of Event Horizons

Robin James Spivey

Biological Sciences, Bangor University, Brambell, Deiniol Road, Bangor, Gwynedd, Great Britain
E-mail: y.gofod@gmail.com

A century has elapsed since gravitational waves were predicted. Their recent detection by the LIGO-Virgo collaboration represents another feather in Einstein's cap and attests to the technological ingenuity of experimentalists. However, the news has been portrayed as affirmation of the existence of black holes, objects whose defining characteristics are event horizons. Whilst a gravitational wave chirp is indicative of coalescing bodies and the inferred masses, $29\pm4M_\odot$ and $36\pm5M_\odot$, rule out neutron stars, a prominent yet overlooked feature in the Hanford and Livingston spectrograms points to a curious mass ejection during the merger process. The spectral bifurcations, beyond which down-chirps are clearly discernible, suggest that a considerable quantity of matter spiralled away from the binary system at the height of the merger. Since accretion disks cannot survive until the latter stages of coalescence, a black hole model seems untenable, and Einstein's expectation that black holes can neither form nor ingest matter in a universe of finite age would appear to be upheld. By virtue of general relativity's logical consistency and the fact that gravity propagates at light speed, gravitational collapse must terminate with the formation of pathology-free temporally suspended objects.

1 The black hole controversy

Einstein realised in 1916 that spacetime could mediate the propagation of energy-transporting gravitational waves travelling at light speed [1]. This entirely theoretical deduction was recently confirmed by the LIGO-Virgo collaboration, demonstrating once again the impeccable physical insights of this great scientist. However, the conclusion drawn on the back of this detection, that coalescing black holes triggered the waves [2], directly contradicts Einstein's published stance [3] regarding the outcome of gravitational collapse.

The first static solution to the field equations of general relativity was found that same year describing the gravitational influence of an idealised, infinite density point mass on asymptotically flat space [4]. Due to the Birkhoff theorem, regions of Schwarzschild's metric accurately represent the gravity external to spherically symmetric bodies such as irrotational stars and planets. However, Einstein appreciated that in the immediate vicinity of Schwarzschild's point mass the solution was physically unrealistic, being unreachable from regions outside the event horizon [3].

Einstein's cogent objection to black holes is easily illustrated by a concrete example. If a ray of light moving directly towards a Schwarzschild black hole can neither arrive at the event horizon nor penetrate it, then no particle can. For a lightlike radial trajectory leading towards the event horizon, the Schwarzschild metric reduces to $(dr/dt)^2 = (2m/r - 1)^2$. Assigning initial coordinates $(r, t) = (r_0, 0)$ to a photon, radius $r_1 < r_0$ is attained at time $t_1 > 0$, which can be readily obtained through integration:

$$t_1 = \int_{r_0}^{r_1} \frac{dr}{2m/r - 1} = \int_{r_0}^{r_1} \left(-1 - \frac{2m}{r - 2m} \right) dr, \qquad (1)$$

$$t_1 = r_0 - r_1 + 2m \ln\left(\frac{r_0 - 2m}{r_1 - 2m} \right). \qquad (2)$$

As the photon nears the horizon, $r_1 \to 2m(1 + \epsilon)$ where $0 < \epsilon \ll 1$. Since ϵ is a factor in the denominator of the logarithm, t_1 grows without limit as $\epsilon \to 0$. Accordingly, even though proper time does not advance for lightlike particles, global relationships within the spacetime impose an insurmountable temporal impediment to their arrival at the event horizon. For timelike particles the situation is much the same. As general relativity is a deterministic theory, this calculation has profound implications, despite its brevity. Even in the most favourable of circumstances a black hole cannot absorb matter and, hence, a universe initially devoid of black holes remains forever devoid of black holes. Since general covariance is integral to this theory, changes of coordinates, as detailed for example in references 5–6 of [2], cannot alter this fundamental conclusion.

Although the stationary black hole metrics satisfy the field equations, they lack a dynamical formation mechanism. It is known that event horizons never quite form during gravitational collapse in a universe of finite age [6–12]. Some theorists claim that infalling matter can arrive at the event horizon of a pre-existing black hole in finite proper time, but in practice this is forbidden by the existence of inviolable temporal relationships that permeate spacetime [13]. In addition, a variety of imprecise arguments commonly advanced for the existence of black holes have been robustly refuted [5].

Some stubborn problems now occupying the time of theoretical physicists are symptomatic of misunderstandings. Belief in black holes has given rise to difficulties such as the information paradox [14], loss of causality within rotating black holes, singularities of infinite mass density and the fact

that when matter is trapped within an event horizon, it has no means of influencing external matter, even gravitationally. Furthermore, tension has arisen between the observed characteristics of certain astrophysical phenomena and popular black hole models. In particular, the finite lifetimes and extreme energetics of quasars and active galactic nuclei (AGN) are difficult to reconcile with nearby galaxy clusters which have only reprocessed around 10% of their primordial gas reserves, yet harbour quiescent galactic nuclei.

The ultrarelativistic emission of charged particles by quasars along biaxial jets alludes to an electromagnetically active central engine of some form. Whereas any charge accruing on a spheroidal black hole would be rapidly neutralised, a gravitationally collapsed object of toroidal topology would be defended by a magnetosphere whose flux lines run locally parallel to its surface [13, 15]. This inference clashes with the "principle of topological censorship", a theorem that is irrelevant if a spacetime has no trapped surfaces [16]. Hence, the characteristics of quasars and AGN offer empirical evidence that gravitational collapse produces "dark holes" lacking event horizons [13]. Quasar extinction would coincide with topological collapse and charge nullification.

Appreciation of the impossibility of event horizon formation inspired the first detailed proposal concerning a future mechanism for dark energy decay. It involves the discharge of vacuum energy via the Unruh effect by intense accelerations exposed within the deepest innards of dark holes [17]. The same work also highlights a novel objection to the existence of black holes relating to their unacceptable influence on the total entropy of the universe. A single supermassive black hole devouring matter could potentially double the entropy of the visible universe in the space of a few seconds, despite poor opportunities for interactions of the captured matter.

2 The dawn of gravitational wave astronomy

Since the announcement that gravitational waves have been detected it has emerged that the GW150914 event closely coincided with a gamma ray burst originating in the same sector of the sky [19]. Suggesting a common source, the binary system must have, as the authors put it, become "unexpectedly active" during coalescence. The possibility that one or more neutron stars were involved can be rejected due to the large masses involved [2,20]. The gamma rays are clearly inconsistent with the no-hair conjecture: any accreting matter should be ejected well before the merger [21]. It is therefore interesting to revisit the gravitational wave data to look for any other evidence of unanticipated peculiarities.

Two such examples draw the eye. In the spectrograms of *both* laser interferometers a down-chirp can be clearly discerned, bifurcating from the somewhat stronger up-chirp during the final crescendo of the merger (see Figure 1). These appear to be comfortably above the noise floor of each detector. The down-chirps are not only present in both spectrograms, they are identically located and share the same characteristics: important hallmarks of a genuine signal.

For a binary dark hole or binary frozen star model, significant mass loss is conceivable during a cataclysmic merger of this kind. The particles held in suspension by time dilation would be strongly perturbed by the gravitational ripples, transporting here a total energy estimated at $3M_{\odot}c^2$ [2]. The combination of this disruption and the violent rotation, particularly during the non-axisymmetric dumbbell phase of coalescence, could plausibly give rise to significant expulsion of matter at the peripheral fringes of the system. The spectral traces are consistent with matter being centrifugally launched with a radial velocity component of approximately $0.04c$. There is also a marked acceleration of the chirp following the shedding of mass, as might be anticipated if the rest mass energy of the ejecta was comparable to the energy radiated in gravitational waves.

From (2), at late times an infalling photon asymptotically approaches the radius $r = 2m$. Why must the photon halt at the exact radius of the event horizon? Why does general relativity only marginally forbid the growth and formation of black holes? Could matters have been any different?

Einstein's theory of gravitation was built upon special relativity which insists that nothing can travel faster than the speed of light in vacuum, prohibiting objects from exerting any form of superluminal influence. As in Newton's theory, gravity has infinite range. This demands that gravitons be massless, with current experimental constraints providing an upper limit of 1.2×10^{-22} eV. Signals from LIGO's geographically separated interferometers support the expectation that gravity travels at the speed of light [2]. Were the speed of gravity any different, the terminal radius of the photon would change, and philosophical problems would ensue.

If photons could only asymptotically approach some radius $r > 2m$, gravitational time dilation could then grow without limit in relatively moderate circumstances, curbing the maximum curvature of spacetime irrespective of Planck-scale limitations. If photons could asymptotically approach some radius $r < 2m$ then event horizons could form, bringing with them all the pathologies associated with black holes. Only if gravity travels at the speed of light can spacetime be arbitrarily warped without fear of event horizon formation, points of infinite mass density, time travel paradoxes or violation of unitarity. Like gravity, electromagnetism has unlimited range. Electric fields are mediated by virtual photons. If black holes did exist then the electric fields of charged particles would vanish upon capture, creating an 'electrical paradox' akin to the very widely acknowledged information paradox. Fortunately, Einstein appears to have formulated a consistent theory of gravitation in which anomalies are avoided but all else is permitted. A strongly curved spacetime may be vital for the timely decay of dark energy [17], a possible requirement for gravity to propagate no slower than light.

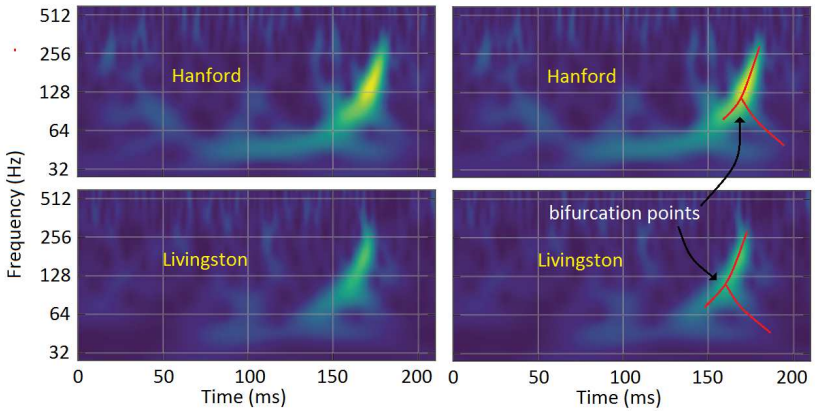

Fig. 1: The gravitational wave spectrograms for the Hanford (top) and Livingston (bottom) Advanced LIGO detectors [2]. Right column: spectral traces have been annotated to show the primary up-chirp and a matching pair of bifurcations beyond which the decline in frequency and amplitude suggests the ejection of mass spiralling away from the merging binary system.

3 Discussion

Gravitational waves have the capability to rectify some long-standing theoretical misconceptions. With improvements in sensitivity already scheduled we shall soon know whether mass-ejections are a generic feature of dark hole coalescence events. If so, we might in time witness some spectacular mergers of supermassive dark holes in the aftermath of galactic mergers within galaxy clusters. For coalescing bodies of large and favourably aligned angular momenta, the resulting gravitational wave signatures could be morphologically very distinct from GW150914 due to the formation of a toroidal dark hole with an unusually lengthy ringdown phase [15, 17]. The publicity and interest surrounding the announcement that gravitational waves have been detected is understandable. However, there has been little or no mention of the fact that the presence of a black hole event horizon cannot be verified even in principle [22] or that Einstein had mathematical grounds for dismissing the notion that black holes exist [3]. Black hole proponents might care to take note that our civilisation still awaits evidence that any of Einstein's predictions concerning gravity were incorrect.

Submitted on February 23, 2016 / Accepted on February 28, 2016

References

1. Einstein A. Approximative integration of the field equations of gravitation. *Sitzungsber. Preuss. Akad. Wiss. Berlin*, 1916, v. 1, 688.
2. Abbott B. P. et al., LIGO & VIRGO collaborations. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 2016, v. 116, 061102.
3. Einstein A. On a stationary system with spherical symmetry consisting of many gravitating masses. *Annals of Mathematics*, 1939, v. 40, 4.
4. Schwarzschild K. Über das gravitationsfeld eines massenpunktes nach der Einsteinschen theorie. *Sitzungsber. Preuss. Akad. Wiss. Berlin*, 1916, v. 1, 189.
5. Weller D. Five fallacies used to link black holes to Einstein's relativistic space-time. *Prog. Phys.*, 2011, v. 7 (1), 93.
6. Oppenheimer J. R. & Snyder H. On continued gravitational contraction. *Phys. Rev.*, 1939, v. 56 (5), 455.
7. Vachaspati T., Stojkovic D. & Krauss L. M. Observation of incipient black holes and the information loss problem. *Phys. Rev. D*, 2007, v. 76 (2), 024005.
8. Kiselev V. V., Logunov A. A. & Mestvirishvili M. A. Black holes: theoretical prediction or fantasy? *Physics of Particles and Nuclei*, 2006, v. 37 (3), 317–320.
9. Kiselev, V. V. E., Logunov A. A. & Mestvirishvili M. A. The physical inconsistency of the Schwarzschild and Kerr solutions. *Theor. & Math. Physics*, 2010, v. 164 (1), 972–975.
10. Shiekh A. Y. Approaching the event horizon of a black hole. *Adv. Stud. Theor. Phys.*, 2012, v. 6 (23), 1147–1152.
11. Chafin C. E. Globally Causal Solutions for Gravitational Collapse. arXiv: gr-qc/1402.1524.
12. Piñol M. A model of dust-like spherically symmetric gravitational collapse without event horizon formation. *Prog. Phys.*, 2015, v. 11 (4), 331.
13. Spivey R. J. Dispelling black hole pathologies through theory and observation. *Prog. Phys.*, 2015, v. 11, 321–329.
14. Mathur S. D. The information paradox: a pedagogical introduction. *Class. Quant. Grav.*, 2009, v. 26 (22), 224001.
15. Spivey R. J. Quasars: a supermassive rotating toroidal black hole interpretation. *MNRAS*, 2000, v. 316 (4), 856–874.
16. Friedman J. L., Schleich K. & Witt D. M. Topological censorship. *Phys. Rev. Lett.*, 1993, v. 71 (10), 1486.
17. Spivey R. J. A non-anthropic solution to the cosmological constant problem. *Prog. Phys.*, 2016, v. 12 (1), 72–84.
18. Spivey R. J. A cosmological hypothesis potentially resolving the mystery of extraterrestrial silence with falsifiable implications for neutrinos. *Physics Essays*, 2015, v. 28 (2), 254–264.
19. Connaughton V. et al. Fermi GBM observations of LIGO gravitational wave event GW150914. arXiv: astro-ph.HE/1602.03920v3.
20. Oppenheimer J. R. & Volkoff G. M. On massive neutron cores. *Physical Review*, 1939, v. 55 (4), 374.
21. Loeb A. Electromagnetic counterparts to black hole mergers detected by LIGO. arXiv: astro-ph.HE/1602.04735v1.
22. Abramowicz M. A., Kluzniak W. & Lasota J. P. No observational proof of the black-hole event-horizon. *Astronomy & Astrophysics*, 2002, v. 396 (3), L31–L34.

Repulsive Gravity in the Oppenheimer-Snyder Collapsar

Trevor W. Marshall
Buckingham Centre for Astrobiology, The University of Buckingham, Buckingham MK18 1EG, UK
E-mail: trevnat@talktalk.net

The Oppenheimer-Snyder metric for a collapsing dust ball has a well defined equilibrium state when the time coordinate goes to plus infinity. The entire ball is contained within the gravitational radius r_0, but half of its content lies within a thin shell between r_0 and $0.94r_0$. This state has the acausal property that no light ray escapes from it, but if one boundary condition at the surface, which Oppenheimer and Snyder imposed without justification, is removed, then all points in the interior remain in causal contact by null geodesics with the exterior. This modification causes the half shell's interior radius to increase to $0.97r_0$. Together with the results of a previous article on the density inside a spherosymmetric neutron star, the present results indicate that, in contrast with the universal attraction of Newtonian gravity, General Relativity gives gravitational repulsion at high density.

1 Introduction

The modern concept of black hole originates with Chandrasekhar's [1] discovery of an upper bound for the mass of a Newtonian white dwarf; it has been claimed (see, for example [2] section 11.3) that the replacement of Newtonian gravitation by General Relativity (GR) makes no significant difference. Using GR, Oppenheimer and Volkoff [3] (OV) found a similar result for neutron stars, the upper bound being somewhat lower than in the white-dwarf case. The OV article, in its footnote 10, did indicate that the GR field equations allow for a stable solution having zero density at the origin in place of the maximum density there of the Newtonian solution, but gave no further attention to this possibility; there seems to have been no serious attempt to return to it since, though a well known text ([5] after equation 23.20) has described it as "unphysical". We showed [4] that solutions of the OV-footnote variety may easily be obtained. The only new feature of such solutions which could conceivably qualify for the "unphysical" label is that the metric has a simple-pole singularity at the origin. This singularity is curiously similar to that now very widely used to describe a black hole, but with the crucial difference that its residue is positive, so that instead of infinite density there we find zero density.

In our previous article we advocated a field, rather than the geometric interpretation of GR, constructing a field energy tensor to explain why the stellar material is concentrated in a spherical shell and not at the origin. Here we shall use an exclusively geometric description, but will nevetheless be able to demonstrate, by studying the particle geodesics inside the shell, that the picture which emerges almost demands that we accept there is gravitational repulsion in the interior of the shell. We conclude that the black hole is a Newtonian concept, superseded by GR.

Our geometric investigation is based on what seems to be the only time-dependent study of a collapsar, namely that of Oppenheimer and Snyder [6] (OS). In an early stage of black-hole theory this article's conclusion was seriously misquoted by Penrose [7] who stated:

"The general situation with regard to a spherically symmetrical body is well known [6]. For a sufficiently great mass, there is *no final equilibrium state* (our emphasis). When sufficient thermal energy has been radiated away, the body contracts and continues to contract until a physical singularity is encountered at $r = 0$."

OS did not say anything resembling this assertion of Penrose. Indeed we shall show below that the OS density distribution approaches a stationary distribution, whose diameter is twice the gravitational radius, as the time goes to plus infinity. It is true that OS also found that in this limit there is a region inside the collapsar from which light may not be emitted, but we shall show below that this is not a real property of the model, and that it may be easily repaired so that all points of the physical space, exterior and interior, remain causally connected at all times. Nobody has demonstrated that any real collapse situation leads to the "trapped surfaces" of the Penrose article, and I would argue that such surfaces would violate the kind of causality described in Weinberg's text ([2] section 7.5). This conclusion was also stated recently by Chafin [8].

2 The OS metric

OS used the comoving coordinates (τ, R, θ, ϕ) with the metric

$$ds^2 = d\tau^2 - \frac{8m^3R}{r}\,dR^2 - r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right), \qquad (1)$$

$$r = 2m\left(R^{3/2} - \frac{3\tau}{4m}\right)^{2/3},$$

in the exterior region $R > 1$ and

$$ds^2 = d\tau^2 - \frac{r^2}{R^2}\,dR^2 - r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right), \qquad (2)$$

$$r = 2mR\left(1 - \frac{3\tau}{4m}\right)^{2/3},$$

in the interior region $0 < R < 1$. By the transformation

$$t = \frac{4m}{3} R^{3/2} - \frac{2}{3}\sqrt{\frac{r^3}{2m}} - 2\sqrt{2mr} + 2m \ln \frac{\sqrt{r} + \sqrt{2m}}{\sqrt{r} - \sqrt{2m}}, \quad (3)$$

the exterior metric converts to the Schwarzschild form

$$ds^2 = \frac{r - 2m}{r} dt^2 - \frac{r}{r - 2m} dr^2 - r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right). \quad (4)$$

We note that, since R is a comoving coordinate, $R = const.$ is a freefall geodesic and in particular the surface $r_1(t)$, that is $R = 1$, satisfies

$$t = \frac{4m}{3} - \frac{2}{3}\sqrt{\frac{r_1^3}{2m}} - 2\sqrt{2mr_1} + 2m \ln \frac{\sqrt{r_1} + \sqrt{2m}}{\sqrt{r_1} - \sqrt{2m}}, \quad (5)$$

and also any such geodesic, for $R > 1$, has its speed v increasing up to a maximum $v = 2c/(3\sqrt{3})$ and then decreasing asymptotically to zero as r approaches $2m$. This confirms the OS statement [6] "...an external observer sees the star asymptotically shrinking to its gravitational radius".

For $0 < R < 1$, OS identified an "internal time" t by defining a *cotime* y as

$$t = \frac{4m}{3} - \frac{4m}{3}\sqrt{y^3} - 4m\sqrt{y} + 2m \ln \frac{\sqrt{y} + 1}{\sqrt{y} - 1}, \quad (6)$$

and then putting

$$y = \frac{r}{2mR} + \frac{R^2 - 1}{2}. \quad (7)$$

This not only gives a continuous match for the internal and external t at $R = 1$, but also the metric in $0 < R < 1$ is

$$ds^2 = \frac{2mr^2(y-1)^2}{Ry^3(r - 2mR^3)} dt^2 - \frac{r}{r - 2mR^3} dr^2 - \\ - r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right), \quad (8)$$

which is continuous with (4) at $R = 1$.

From (6) and (7) we now see that the equilibrium state of the OS model is given by

$$r = mR(3 - R^2) \quad (0 < R < 1), \quad (9)$$

which contradicts the conclusion stated by Penrose and quoted in the previous section of this article. The density ρ is obtained from the curvature tensor of (1)

$$\rho = \frac{mR^3}{4\pi r^3}, \quad (10)$$

and since

$$\sqrt{-g} = \frac{r^3}{R}\sin\theta, \quad (11)$$

it integrates over the volume of the collapsar to give

$$\int_{R<1} \rho\sqrt{-g}\, dR\, d\theta\, d\phi = m. \quad (12)$$

In the remote past, when $r \sim yR, y \to \infty$, the dust particles are distributed uniformly over the sphere's interior, but as collapse proceeds their trajectories, $R = const$, crowd near the surface. This may be shown by considering that in the remote past half of the particles are contained within a shell between $R = 2^{-1/3} = 0.7937$ and $R = 1$, and that their final positions are $r = mR(3 - R^2)$, so that they end up between $r = 1.881m$ and $r = 2m$.

3 A problem with causality

At no time does the entire content of the collapsar go inside the sphere $r = 2m$, so Figure 1 of Penrose [7] is an incorrect picture of the OS collapsar, as is the discussion about *trapped surfaces* on which the figure is based. There is, however a causal anomaly in the OS model, in that, for any $R < 1$, there is a value of t beyond which no light signal emerges.

For the region $R < 1$ we introduce the coordinates (x, R, θ, ϕ), where $x = r/(2mR)$. The metric is

$$\frac{1}{4m^2} ds^2 = xdx^2 - x^2 dR^2 - x^2R^2\left(d\theta^2 + \sin^2\theta d\phi^2\right). \quad (13)$$

A radial light wave or radial null geodesic (RNG) satisfies

$$\frac{dR}{dx} = -\frac{1}{\sqrt{x}}, \quad (14)$$

that is

$$R = 2\sqrt{x(0)} - 2\sqrt{x} = 2\sqrt{x_0} - 2\sqrt{x}. \quad (15)$$

In order to reach the surface at $R = 1$ we need $x(1) > 1$ and therefore $x_0 > 9/4$, but from (7) we find that the minimum value of x_0 is 3/2, reached at cotime $y = 1$, that is when t is plus infinity. It follows that, for $y > 7/4$, an RNG from the origin cannot escape.

It is a simple matter to repair this flaw in the OS model; we replace (7) by

$$y = 1 + x - \frac{(R-3)^2}{4}, \quad (16)$$

so that all RNGs for $y > 1$ escape and causality is preserved. I have established [9] that the metric tensor with these coordinates is again continuous at $R = 1$. It differs from (8) in that the tensor component g_{rt} is not zero in $R < 1$, but only at $R = 1$. It was the unjustified imposition of the condition $g_{rt} = 0$ which led OS to claim that the connection (7) between the (x, R) and the (t, r) coordinates is unique. Our amendment of the OS metric leads to a more concentrated shell, because the equilibrium state is now specified by

$$r = \frac{mR(3-R)^2}{2}, \quad (17)$$

which, putting $R = 2^{-1/3}$, leaves half of the original dust matter in a shell between $r = 1.932m$ and $r = 2m$.

4 Gravity becomes repulsive at high densities

We have all believed since 1687 that gravity is universally attractive, so it requires some effort to adjust to the idea that gravity may repel; even the new mode of thought which came with GR did not change the paradigm of attractive gravity. We have attempted to show elsewhere [4] how a full appreciation of the gravitational field may cause us to change our intuition. However, for the present article we shall stay within the geometric presentation of GR, merely pointing the way towards an understanding of repulsive gravity.

We consider the motion of a foreign dust particle of small mass which crashes radially into the surface $R = 1$ at time t, that is at the point $r = r_1(t)$ given by (5), with a speed greater than that at which the surface itself is moving. We ignore the gravitational force exerted by this foreign particle, so it moves along a radial geodesic of the metric (13). The coordinate R is cyclic, so we have a conservation equation

$$x^2 \frac{dR}{ds} = -C\,, \quad (C > 0)\,, \tag{18}$$

and it then follows that

$$\frac{dR}{dx} = \frac{C}{\sqrt{C^2 x + x^3}}\,. \tag{19}$$

This equation, when integrated with initial conditions $(x, R) = (r_1/2m, 1)$, leads to a relation between the final values x_∞ and R_∞ at t equal to plus infinity

$$R_\infty = 1 - \int_{x_\infty}^{r_1/2m} \frac{C}{\sqrt{C^2 x + x^3}}\, dx\,. \tag{20}$$

Now substituting $y = 1$ in (16) provides a second such relation, so eliminating x_∞ we obtain R_∞ in terms of r_1 and C. This is not a difficult process numerically, but in the limiting ultrarelativistic case $C \to \infty$ – effectively a null geodesic – it becomes especially simple

$$R_\infty = 2 - \sqrt{r_1/(2m)}\,, \quad (r_1 < 8m)\,. \tag{21}$$

If $r_1 > 8m$ such a particle passes through the centre and exits at the opposite end of the diameter. A particle which crashes into the collapsar when the latter is close to its final state – r_1 close to 1 – does not penetrate it beyond the surface shell described in the previous section.

As long as we stay within the constraints of the geometric interpretation of GR, we are not able to draw inferences about what causes such a dramatic deceleration; we could, for example [10], continue to insist that it results from time dilation of the metric. I suggest, however, that a return to the language of field theory offers us, at the very least, an attractive alternative; we may claim that the force of repulsive gravity which decelerates the incident particle is the very same as the one which compresses the particles of the collapsar into a thin shell. In the context of a collapsar having a more realistic equation of state we pursued this point of view in our previous article [4].

Submitted on March 2, 2016 / Accepted on March 6, 2016

References

1. Chandrasekhar S. Stellar Structure. Dover, New York, 1939, Chapter IV.
2. Weinberg S. Gravitation and Cosmology. John Wiley, New York, 1972.
3. Oppenheimer J. R. and Volkoff G. M. *Phys. Rev.*, 1939, v. 54, 540.
4. Wallis M. K. and Marshall T. W. Energy in General Relativity — the case of the neutron star. *Proc. Physical Interpretations of Relativity Theory*, 2015, Bauman State Technical University, Moscow, 544–556.
5. Misner C. W., Thorne K. S. and Wheeler J. A. Gravitation. Freeman, San Francisco, 1970.
6. Oppenheimer J. R. and Snyder H. *Phys. Rev.*, 1939, v. 56, 455.
7. Penrose R. *Phys. Rev. Lett.*, 1965, v. 14, 57.
8. Chafin C. E. Globally Causal Solutions for Gravitational Collapse. arxiv: gr-qc/1402.1524.
9. Marshall T. W. *Astrophys. Space Sci.*, 2012, v. 342, 329–332.
10. Thorne K. S. Black Holes and Time Warps. Norton, New York, 1994, p. 417.

The Dirac-Electron Vacuum Wave

William C. Daywitt

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

This paper argues that the Dirac equation can be interpreted as an interaction between the electron core and the Planck vacuum state, where the positive and negative solutions represent respectively the dynamics of the electron core and a vacuum wave propagating within the vacuum state. Results show that the nonrelativistic positive solution reduces to the Schrödinger wave equation.

1 Introduction

In its rest frame the massive electron core $(-e_*, m)$ exerts the two-term coupling force [1, Sec.7-8]

$$F(r) = \frac{e_*^2}{r^2} - \frac{mc^2}{r} = \frac{(-e_*)(-e_*)}{r^2} - \frac{mm_*G}{r_* r} \qquad (1)$$

on the PV quasi-continuum, where e_* is the massless bare charge with its derived electron mass m, and $G\,(= e_*^2/m_*^2)$ is Newton's gravitational constant. The first $(-e_*)$ in (1) belongs to the electron and the second to the separate Planck particles making up the degenerate PV state. The two terms in (1) represent respectively the Coulomb repulsion between the electron charge and the separate PV charges, and their mutual gravitational attraction.

The particle/PV coupling force (1) vanishes at the electron Compton radius $r_c\,(= e_*^2/mc^2)$. In addition, the vanishing of $F(r_c)$ is a Lorentz invariant constant [2] that leads to the important Compton-(de Broglie) relations

$$r_c \cdot mc^2 = r_d \cdot cp = r_L \cdot E = r_* \cdot m_* c^2 = e_*^2 \;(= c\hbar) \qquad (2)$$

where $r_d = r_c/\beta_0\gamma_0$ and $r_L = r_c/\gamma_0$, and $r_*\,(= e_*^2/m_*c^2)$ and m_* are the Compton radius and mass of the Planck particles within the PV. The ratio of the electron speed v to the speed of light c is β_0 and $\gamma_0 = 1/(1-\beta_0^2)^{1/2}$. The relativistic momentum and energy following from the invariance of $F(r_c) = 0$ are $p\,(= m\gamma_0 v)$ and $E\,(= m\gamma_0 c^2)$, from which $E = (m^2c^4 + c^2p^2)^{1/2}$ is the relativistically important energy-momentum relationship.

The results of the previous paragraph show that the important relativistic energy E and momentum p (or its vector counterpart $\mathbf{p}$) are determined at the basic core-PV interaction level. Furthermore, since the core is many orders-of-magnitude smaller than the electron Compton radius, it is reasonable to assume that this point-core picks up its wave-particle nature (including its Compton radius and its energy and momentum operators) from its coupling to the PV continuum.

2 Dirac equation

The Dirac equation [3, p.79]

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\phi\\ \chi\end{pmatrix} = \begin{pmatrix}c\vec{\sigma}\cdot\widehat{\mathbf{p}}\chi\\ c\vec{\sigma}\cdot\widehat{\mathbf{p}}\phi\end{pmatrix} + mc^2\begin{pmatrix}\phi\\ -\chi\end{pmatrix} \qquad (3)$$

where $\widehat{\mathbf{p}}\;(=-i\hbar\nabla)$ is the momentum operator and $\hbar$ is the reduced Planck constant, can be expressed using (2) as

$$ir_c\frac{\partial}{c\partial t}\begin{pmatrix}\phi\\ \chi\end{pmatrix} + \begin{pmatrix}\vec{\sigma}\cdot ir_c\nabla\chi\\ \vec{\sigma}\cdot ir_c\nabla\phi\end{pmatrix} = \begin{pmatrix}\phi\\ -\chi\end{pmatrix} \qquad (4)$$

where the solutions ϕ and χ for this electron-vacuum system are 2x1 Dirac spinors, and $\vec{\sigma}$ is the Pauli 2x2 vector matrix derived from the three 2x2 Pauli spin matrices σ_k $(k = 1, 2, 3)$ [3, p.12]

$$\sigma_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad \sigma_2 = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\quad \sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \qquad (5)$$

The gradient operators $\partial/c\partial t$ and ∇ in (4) are normalized by the electron Compton radius r_c

$$\frac{\partial}{c\partial t/r_c} \qquad \text{and} \qquad \frac{\partial}{\partial x_k/r_c} \qquad (6)$$

whose denominators can be looked upon as the normalized line elements cdt/r_c and dx/r_c of a spacetime [4, p.27] perturbed by the electron core $(-e_*, m)$ in (1). Following from this viewpoint is the concept that *the 2x1 spinors ϕ and χ represent, respectively, the PV response to the electron core $(-e_*, m)$ and some type of vacuum wave.* Furthermore, the vacuum wave cannot be a Planck-particle wave, since the PV is a degenerate state (where the vacuum eigenstates are fully occupied). Thus the wave must be of the nature of a percussion wave, analogous to a wave traveling on the head of a kettle drum.

3 Dirac-Schrödinger reduction

The solution χ in the two simultaneous equations of (4) is assumed in the PV theory to represent a relativistic vacuum wave propagating within the PV state. What follows derives the nonrelativistic version of that wave to add more credence and understanding the vacuum wave idea.

The Dirac-to-Schrödinger reduction [3, p.79] of (4) begins with the elimination of its mass related, high-frequency, components by assuming

$$\begin{pmatrix}\phi\\ \chi\end{pmatrix} = \begin{pmatrix}\phi_0\\ \chi_0\end{pmatrix}e^{-i(mc^2t/\hbar)} = \begin{pmatrix}\phi_0\\ \chi_0\end{pmatrix}e^{-i(ct/r_c)} \qquad (7)$$

where ϕ_0 and χ_0 are slowly varying functions of time compared to the exponential. This result implies that the frequency $\omega_c = c/r_c \gg \omega_0$ for any ω_0 associated with ϕ_0 or χ_0. Inserting (7) into (4) gives

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix}\phi_0\\ \chi_0\end{pmatrix} + \begin{pmatrix}\overrightarrow{\sigma}\cdot ir_c\nabla\chi_0\\ \overrightarrow{\sigma}\cdot ir_c\nabla\phi_0\end{pmatrix} = \begin{pmatrix}0\\ -2\chi_0\end{pmatrix} \tag{8}$$

where the 0 on the right is a 2x1 null spinor. This zero spinor indicates that the mass energy of the free electron core is being ignored, while the effective negative mass-energy of the vacuum wave has doubled ($-2\chi_0$). In effect, mass energy for the core-vacuum system has been conserved by shifting the mass energy from the free relativistic core to the vacuum wave.

The lower of the two simultaneous equations in (8) can be reduced from three to two terms by the assumption

$$\left| ir_c \frac{\partial\chi_0}{c\partial t} \right| \ll |-2\chi_0| \tag{9}$$

if the kinetic energy of the vacuum wave is significantly less than its effective mass energy. Inserting (9) into (8) then yields

$$ir_c \frac{\partial}{c\partial t}\begin{pmatrix}\phi_0\\ 0\end{pmatrix} + \begin{pmatrix}\overrightarrow{\sigma}\cdot ir_c\nabla\chi_0\\ \overrightarrow{\sigma}\cdot ir_c\nabla\phi_0\end{pmatrix} = \begin{pmatrix}0\\ -2\chi_0\end{pmatrix} \tag{10}$$

as the nonrelativistic version of (4). The mass energy of the free core, and the kinetic energy of the vacuum wave (associated with the lower-left null spinor), are discarded in this nonrelativistic approximation to (4).

Separating the two equations in (10) produces

$$ir_c \frac{\partial\phi_0}{c\partial t} + \overrightarrow{\sigma}\cdot ir_c\nabla\chi_0 = 0 \tag{11}$$

and

$$\overrightarrow{\sigma}\cdot ir_c\nabla\phi_0 = -2\chi_0 \tag{12}$$

where the second term in (11) and the first term in (12) represent the connection between the free-space core dynamics (ϕ_0) and the vacuum wave (χ_0). Inserting (12) into (11) then leads to [3, p.80]

$$ir_c \frac{\partial\phi_0}{c\partial t} - \frac{(\overrightarrow{\sigma}\cdot ir_c\nabla)^2}{2}\phi_0 = 0\,. \tag{13}$$

Finally, using the Pauli-matrix identity [3, p.12]

$$(\overrightarrow{\sigma}\cdot\nabla)^2 = I\,(\nabla)^2 \tag{14}$$

in (13) yields the free-core Schrödinger equation

$$ir_c \frac{\partial\phi_0}{c\partial t} = \frac{(ir_c\nabla)^2}{2}\phi_0 \quad \text{or} \quad i\hbar\frac{\partial\phi_0}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\phi_0 \tag{15}$$

where the two spin components in ϕ_0 are ignored in this nonrelativistic approximation; so ϕ_0 becomes a simple scaler wavefunction rather than a 2x1 spinor.

4 Conclusions and comments

Although the spin components are missing from the standard version of the Schrödinger equation [5, p.20], the solutions to (11) and (12) indicate that those components are still meaningful.

Using $r_c\,(= e_*^2/mc^2 = \hbar/mc)$ from (2) in (11) and (12) yields

$$ir_c \frac{\partial\phi_0}{c\partial t} = \overrightarrow{\sigma}\cdot(\widehat{\mathbf{p}}/mc)\,\chi_0 \tag{16}$$

and

$$\chi_0 = \overrightarrow{\sigma}\cdot(\widehat{\mathbf{p}}/2)\,\phi_0 \tag{17}$$

where $\widehat{\mathbf{p}}\,(= -i\hbar\nabla)$ is the vector momentum operator.

Equations (16) and (17) from the perturbed spacetime can be understood as follows: the free-space energy from ϕ_0 in the first term of (16) drives the vacuum energy associated with the second term; this χ_0 energy of the second term in (17) then feeds back into the ϕ_0 term in (16), leading to a circular simultaneity between the two equations that represent the coupled nonrelativistic behavior of the core-PV system. Furthermore, the fact that there is no kinetic-energy term in (17) suggests that the localized energy in the PV travels as a percussion wave through that vacuum state. This scenario represents the PV view of the Dirac electron equation (4): that is, the dynamics of the free-space electron core $(-e_*, m)$ lead to a vacuum wave propagating within the PV state, in step with the free electron core.

Dedication

This paper is dedicated to the memory Dr. Petr Beckmann [6], Professor Emeritus Electrical Engineering, the University of Colorado, Boulder, Colorado.

Submitted on March 18, 2016 / Accepted on March 20, 2016

References

1. 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 (1–1), 22. See also www:planckvacuum.com.
2. Daywitt W. C. The de Broglie relations derived from the electron and proton coupling to the Planck vacuum state. *Progress in Physics*, 2015, v. 11 (2), 189.
3. Gingrich D.M. Practical Quantum Electrodynamics. CRC, The Taylor & Francis Group, Boca Raton, 2006.
4. Leighton R.B. Principles of Modern Physics. McGraw-Hill Book Co., New York, 1959.
5. Schiff L.I. Quantum Mechanics, 2nd. McGraw Book Co., Inc., New York, 1955.
6. Petr Beckmann. Article from Wikipedia.

Mass of a Charged Particle with Complex Structure in Zeropoint Field

Kundeti Muralidhar

Physics Department, National Defence Academy, Khadakwasla, Pune-411023, India. kundetimuralidhar@gmail.com

A charged particle immersed in the fluctuating zeropoint field may be visualized as an oscillator and such an oscillating particle is considered to possess an extended structure with center of mass and center of charge separated by radius of rotation in a complex vector space. Considering stochastic electrodynamics with spin, the zeropoint energy absorbed by the particle due to its internal motion has been derived. One may initially assume a massless charged particle with complex structure and after interaction with zeropoint field, the absorbed energy of the particle may correspond to the particle mass. This gives an idea that an elementary particle may acquire mass from the interaction of zeropoint field. When the particle moves as a whole, there appears to be a small energy correction of the order of fine structure constant and it may be attributed to the mass correction due to particle motion in the zeropoint field.

1 Introduction

The Dirac electron executes rapid oscillations superimposed on its normal average translational motion and this oscillatory motion is known as zitterbewegung and it was first shown by Schrödinger. In the zitterbewegung motion, the electron appears vibrating rapidly with a very high frequency equal to $2mc^2\hbar^{-1}$ and with internal velocity equal to the velocity of light. These oscillations are confined to a region of the order of Compton wavelength of the particle. It has been shown by several authors over decades that the center of charge and center of mass of charged particle are not one and the same but they are separated by a distance of the order of Compton wavelength of the particle. The approach of extended particle structure was developed by Wyssenhoff and Raabe [1], Barut and Zhanghi [2], Salesi and Recami [3] and others. The list of references connected with the validity of the extended or internal structure of charged particle are too many and some of them are mentioned in the reference [4]. Thus the structure of an elementary charged particle is not definitely a point particle with charge and mass or a spherical rigid body with charge distribution. The structure of electron may be visualized as the point charge in a circular motion with spin angular momentum. The frequency of rotation is equal to the zitterbewegung frequency and the radius of rotation is equal to half the average Compton wavelength. The circular motion is observed from the rest frame positioned at the centre of rotation which is the centre of mass point. Thus the centre of mass point and the centre of charge point are separated by the radius of rotation. The electron spin generated from the circular motion of zitterbewegung was advocated by several researchers. Holten [5] discussed the classical and quantum electrodynamics of spinning particles. In the Holten theory, the spinning particle emerges as a modification of relativistic time dilation by a spin dependent term and the zitterbewegung appears as a circular motion and the angular momentum of such circular motion represents the spin. In the Hestene model of Dirac electron [6], the spin was considered as a dynamical property of the electron motion. In the approach of geometric algebra, using multivector valued Lagrangian, the angular momentum of this internal rotation represents particle spin and it has been explicitly shown as a bivector quantity representing the orientation of the plane of rotation [7, 8]. In quantum theories, the internal oscillations of the particle are attributed due to vacuum fluctuations. However, in stochastic electrodynamics, the internal oscillatory motion of the particle is attributed to the presence of zeropoint field throughout space [9]. The mass of the particle is seen as the energy of oscillations confined to a region of space of dimensions of the order of Compton wavelength [10].

The classical concept of space is an infinite void and featureless. However, it has been replaced by the vacuum field or the zeropoint random electromagnetic field when the quantum oscillator energy was found to contain certain zeropoint energy and with the substitution of the quantum oscillator energy into the Planck's radiation formula yields the energy density of zeropoint field at absolute zero temperature [11]. In a classical approach to the radiation problem, Einstein and Stern obtained blackbody radiation spectrum and suggested that a dipole oscillator possessed zeropoint energy. In 1916, Nernst proposed that the universe might actually contain ubiquitous zeropoint field without any presence of external electromagnetic sources [12, 13]. Thus the origin of zeropoint field is presumed to be purely a quantum mechanical effect and considered to be uniformly present throughout space in the form of stochastic fluctuating electromagnetic field. The zeropoint radiation is found to be homogeneous and isotropic in space. The spectral density of zeropoint radiation is proportional to ω^3 and it is therefore Lorentz invariant. The electromagnetic zeropoint field consists of fluctuating radiation that can be expressed as a superposition of polarised plane waves. Because of the random impulses from fluctuating zeropoint field, a free particle cannot remain at rest but oscillates about its equilibrium position.

The Planck's idea of zeropoint radiation field was revis-

ited by Marshall and explicitly showed that the equivalence between classical and quantum oscillators in the ground state [14]. This has inspired interesting modifications to classical electrodynamics and the developed subject is called stochastic electrodynamics. Stochastic electrodynamics deals with the movement of charged particles in the classical electromagnetic fluctuating zeropint field. The presence of classical, isotropic, homogeneous and Lorentz invariant zeropoint field in the universe is an important constituent of stochastic electrodynamics. The stochastic electrodynamics approach was used to explain classically several important fundamental results and problems of quantum mechanics [15–20]. Boyer [15] showed that for a harmonic oscillator, the fluctuations produced by zeropint field are exactly in agreement with the quantum theory and as a consequence the Heisenberg minimum uncertainty relation is satisfied for the oscillator immersed in the zeropoint field. Stochastic electrodynamics was used to explain the long standing problems of quantum mechanics, namely the stability of an atom, Van der Walls force between molecules [16], Casimir force [17], etc. All these studies reveal the fact that the conventional concept of space has been changed by the emergence of zeropoint field. A detailed account of stochastic electrodynamics as a real classical electromagnetic field and a phenomenological stochastic approach to the fundamental aspects of quantum mechanics was given by de La Pena *et al.,* [13, 21]. In the stochastic electrodynamics, if the upper cut-off frequency to the spectrum of zeropoint field is not imposed, the energy of the oscillator would be divergent. Despite of its success in explaining several quantum phenomena, the results obtained in the stochastic electrodynamics have certain drawbacks [20]; it neglects Lorentz force due to zeropoint magnetic field, it fails in the case of nonlinear forces, explanation of sharp spectral lines is not possible, diffraction of electrons cannot be explained and further the Schrödinger equation can be derived in particular cases only.

A charged point particle immersed in the fluctuating electromagnetic zeropoint field is considered as an oscillator. In the stochastic electrodynamics approach, the equation of motion of the charged particle in the zeropoint field is known as Brafford-Marshall equation [13] which is simply the Abraham-Lorentz [22] equation of motion of a charged particle of mass m and charge e and it is given by

$$m\ddot{\mathbf{x}} - \Gamma_a m\ddot{\mathbf{v}} + m\omega_0^2\mathbf{x} = e\mathbf{E}_z(\mathbf{x}, t), \tag{1}$$

where $\Gamma_a = 2e^2/3mc^3$, ω_0 is the frequency of oscillations of the particle, $\mathbf{v}$ is the velocity of the particle, c is the velocity of light, $\mathbf{E}_z(\mathbf{x}, t)$ is the external electric zeropoint field and an over dot denotes differentiation with respect to time. In the above equation, the force term contains three parts; the binding force $m\omega_0^2\mathbf{x}$, damping force $\Gamma_a m\ddot{\mathbf{v}}$ and external electric zeropoint field force $e\mathbf{E}_z(\mathbf{x}, t)$. In the case of point particles, the strength of these forces follows the relation

$$m\omega_0^2\mathbf{x} < \Gamma_a m\ddot{\mathbf{v}} < e\mathbf{E}_z(\mathbf{x}, t). \tag{2}$$

The energy absorbed by the particle oscillator in the zeropoint field was given by several authors by introducing certain approximations. There are two main approaches found in the literature; one is due to Boyer [6] and the other is due to Rueda [19]. In addition to these main approaches, recently Cavalleri *et al.,* [20] introduced stochastic electrodynamics with spin and explained several interesting phenomena for example, stability of elliptical orbits in an atom, the origin of special relativity and the explanation for diffraction of electrons. It has been shown that the drawbacks of stochastic electrodynamics can be removed with the introduction of spin into the problem. The particle has a natural cut-off frequency equal to the spin frequency which is the maximum frequency radiated by the electron in the zitterbewegung interpretation. This eliminates the problem of divergence in stochastic electrodynamics. These recent advancements in the field of stochastic electrodynamics fully support the assumption that the stochastic electromagnetic field represents the zeropoint field and renew the interest in studying the fundamental aspects of quantum systems and in particular the charged particle oscillator in zeropoint fields.

In Boyer's extensive studies, the harmonic oscillator was developed under the dipole approximation and the charged particle was considered as a point particle without any internal structure. The point particle limit is endowed with two assumptions; i) when the particle size tends to zero, $\omega_c\tau \ll 1$, where ω_c is the cut-off frequency and τ is the characteristic time and ii) when the radiation damping term is very small compared to the external force, $\Gamma_a\omega_c \ll 1$. In Boyer's process of finding the zeropoint energy associated with the charged particle, an integral under narrow line width approximation was solved and finally the zeropoint energy per mode of the oscillator was obtained [16]. This energy has been shown to be equal to the zeropoint energy of the quantum oscillator.

In Rueda's approach, the classical particle was considered as a homogeneously charged rigid sphere and to find the energy absorbed by the particle, the radiation damping and binding terms were neglected when compared to the force term in the Lorentz Abraham equation of motion. The integration was performed over a range 0 to τ, where τ is the characteristic time taken by the electromagnetic wave to traverse a distance equal to the diameter of the particle. The main difference from Boyer's approach is that Rueda assumed $\omega_c\tau \gg 1$ and this condition means the cut-off wavelength is much smaller than the particle size. Further, Rueda introduced a convergence factor $\eta(\omega)$ in the zeropoint energy of the particle oscillator. The average zeropint energy of the oscillator is given by [19]

$$\langle E_0\rangle = \frac{\Gamma_a\hbar\omega_c^2}{\pi}\,\eta(\omega_c). \tag{3}$$

In the later studies, Haitch, Rueda and Puthoff [23] studied an accelerated charged particle under the influence of zeropint field and obtained a relation for inertial mass of a charged particle which is similar to (3). Recently, Haitch *et al.* [9] suggested that the radiation damping constant in the zeropoint field as Γ_z which is not necessarily equal to the damping constant Γ_a of Larmor formula for power radiated by an accelerated charged particle. If we set $\eta(\omega)\Gamma_z\omega_c \sim 1$, the ground state energy of the particle oscillator in the zeropoint field is written as $(\hbar\omega_c)/\pi$. In the case the cut-off frequency is simlar to the resonant frequency of the particle oscillator in the electromagnetic zeropoint field, the ground state energy is equal to the zitterbewegung energy of the Dirac electron. Here, the frequency ω_c is not generally equal to the frequency of oscillation of the particle and it differs by a fraction of fine structure constant. However, the reason for assuming Γ_a as Γ_z is obscure. It may be understood that the energy in (3) corresponds only to a mass correction but not to the mass of the charged particle.

In the stochastic electrodynamics with spin, the particle is considered to possess an extended internal structure and the particle spin is sensitive to the zeropoint frequency that is equal to the frequency of gyration. The particle gyration motion explains the spin properties and refers to a circular motion at the speed of light [20]. The velocity of the particle is not the real velocity of gyrating particle, but centre of mass point around which the particle revolves. The special relativity is not present at the particle level and arises mainly because of the helical motion of the particle when observed from an arbitrary inertial frame of reference [24]. The centre of circular motion responds only to the force parallel to the spin direction. The equation of motion of centre of mass point can be expressed by (1) provided the external force is parallel to the spin direction.

Clifford algebra or Geometric algebra has been considered to be a superior mathematical tool to express many of the physical concepts and proved to provide simpler and straightforward description to the mathematical and physical problems. The geometric algebra was rediscovered by Hestenes [25] in 1960's and it is being used by a growing number of physicists today. In Geometric algebra, a complex vector is defined as a sum of a vector and a bivector. In the complex vector algebra, the oscillations of a charged particle immersed in zeropoint field have been studied recently by the author [26]. The oscillations of the particle in the zeropoint field may be considered as complex rotations in complex vector space. The local particle harmonic oscillator is analysed in the complex vector formalism considering the algebra of complex null vectors. It has been shown that the average zeropoint energy of the particle is proportional to particle bivector spin and the mass of the particle may be interpreted as a local spatial complex rotation in the rest frame.

In the electromagnetic world, the particle mass originates from the electromagnetic field and it is purely electromagnetic in nature [27]. In the classical Lorentz theory of electron, the self-energy is closely connected to the electromagnetic mass of the electron. The self-energy problem in classical theory or quantum theory is essentially connected to the structure of electron and it may not be correct to assign the structure to the electron as a form factor [28]. Further the classical electromagnetic field may be only responsible for the interaction and gives the particle mass as purely electromagnetic in nature. In quantum field theories, the energy, momentum and charge of a particle appear as a consequence of field quantisation and leads to natural classification of particles depending on their spin values. In the renormalization procedure of quantum field theory with finite cut-off for the radiatively induced mass, it has been shown that mass depends on particle spin in the limit when the bare mass tends to zero [29]. However, in the quantum electrodynamics it is well known that the sum of bare mass and the mass correction equals the electron mass and the mass correction is due to the interaction of the particle with vacuum fluctuations [11]. Recently, Pollock interpreted particle mass (fermion or boson) arising from the zeropoint vacuum oscillations by introducing a matrix mass term in the Dirac equation [30]. The standard model deals with the fundamental particles through interaction of bosons, and at a deeper level one may consider the particles as field excitations. Though the vacuum fluctuations have been treated in a different manner in quantum theory and in quantum electrodynamics, the particle oscillations considered either in the vacuum field or in the classical stochastic electrodynamics with spin, are attributed to the fluctuations of the zeropoint field. The idea that the mass arises from the external electromagnetic interaction may lead to the conclusion that charge retains intrinsic masslessness [31]. It has been argued that for there to be correspondence with the particle mass, perhaps at pre-quantum level, inertial mass must originate from external electromagnetic interaction [32].

The aim of this article is to find the energy absorbed by the particle due to its intrinsic motion in the presence of zeropoint field and to discuss the possible origin of mass generation. In section 2, we have explained the modalities of the extended structure of the charged particle in the complex vector algebra. In the present extended particle structure, since we have considered the center of mass point and center of charge separated by radius of rotation in the complex plane, the equation of motion of the particle as a whole is considered as a combination of equation of motion of center of charge and the equation of motion of center of mass. These equations of motion of center of charge and center of mass are derived in section 3. Considering the equation of motion of center of charge in the zeropoint field, the energy absorbed by an extended charged particle is obtained in section 4, and the possible origin of mass generation is discussed in section 5. Finally, conclusions are presented in section 6. Throughout this article a charged particle implies a particle like electron.

2 The complex structure of a charged particle

In the extended particle structure, the centre of mass and the centre of charge positions are considered as separate. Denoting the centre of local complex rotations by the position vector $\mathbf{x}$ and the radius of rotation by the vector ξ , a complex vector connected with both the motion of the centre of mass point and internal complex rotation is expressed as [26]

$$X(t) = \mathbf{x}(t) + \mathbf{i}\,\xi(t). \tag{4}$$

In the geometric algebra, a bivector represents an oriented plane and $\mathbf{i}$ is a pseudoscalar which represents an oriented volume [33]. Differentiating (4) with respect to time gives the velocity complex vector.

$$U(t) = \mathbf{v}(t) + \mathbf{i}\,\mathbf{u}(t). \tag{5}$$

Here, the velocity of centre of mass point is $\mathbf{v}$ and the internal particle velocity is $\mathbf{u}$. A reversion operation on U gives $\bar{U} = \mathbf{v} - \mathbf{i}\,\mathbf{u}$ and the product

$$U\bar{U} = v^2 + u^2. \tag{6}$$

In the particle rest frame $\mathbf{v} = 0$ and $U\bar{U} = u^2$. Since the particle internal velocity in the particle rest frame $u = c$ the velocity of light, $|U| = u = c$. However, when the particle is observed from an arbitrary frame different from the rest frame of the particle centre of mass, as the centre of mass moves with velocity $\mathbf{v}$, the particle motion contains both translational and internal rotational motion of the particle. Then the particle internal velocity can be seen as

$$u^2 = c^2 - v^2 \tag{7}$$

or

$$u = c(1 - \beta^2)^{1/2} = c\gamma^{-1}, \tag{8}$$

where $\boldsymbol{\beta}= \mathbf{v}/c$ and the factor γ is the usual Lorentz factor. The angular frequency of rotation of the particle internal motion is equal to the ratio between the velocity c and radius of rotation ξ, $\omega_s = c/\xi$. When observed from an arbitrary frame, the angular frequency ω would be equal to the ratio between u and ξ

$$\omega = \frac{u}{\xi} = \omega_s\gamma^{-1}. \tag{9}$$

Thus the angular frequency of rotation decreases when observed from an arbitrary frame and the decrease depends on the velocity of the centre of mass. Considering the helical motion of the particle, this method of calculation for time dilation was first shown in a simple manner by Cavelleri [24]. The above analysis shows that the basic reason for the relativistic effects that we observe is due to the internal rotation which is a consequence of fluctuating zeropoint field and elucidates a deeper understanding of relativity at particle level in addition to the constancy of velocity of light postulate. The difference between ω and ω_s corresponds to the particle velocity. In other words, when the particle moves with velocity $\mathbf{v}$, an important consequence is that the particle itself induces certain modification in the field to take place at a lower frequency ω_B. Thus the motion of a free particle is conveniently visualized as a superposition of frequencies ω_0 and ω_B such that the particle motion as observed from an arbitrary frame appears to be a modulated wave containing internal high frequency ω_0 and an envelope frequency ω_B. The ratio between the envelop frequency and the internal frequency is then expressed as

$$\frac{\omega_B}{\omega_0} = \frac{v}{c}. \tag{10}$$

This result is simply a consequence of superposition of internal complex rotations on translational motion of the particle. The relativistic momentum of the center of mass point can be expressed as $\mathbf{p} = \gamma m\mathbf{v}$ and in the complex vector formalism momentum complex vector is given by [26]

$$P = \mathbf{p} + \mathbf{i}\,\pi, \tag{11}$$

where $\pi = m\mathbf{u}$. The total energy of the particle is now expressed as

$$E^2 = P\bar{P}c^2 = (\mathbf{p} + \mathbf{i}\,\pi)(\mathbf{p} - \mathbf{i}\,\pi) = p^2c^2 + m^2c^4. \tag{12}$$

However, in the presence of external electromagnetic field we normally replace the momentum by $\mathbf{p} - e\mathbf{A}/c$ in the minimal coupling prescription. Now, using $\mathbf{p} \rightarrow \mathbf{p} - e\mathbf{A}/c$ in (12) and equating the scalar parts, the total energy of the particle becomes

$$E^2 = p^2c^2 - 2ec\mathbf{p}.\mathbf{A} + e^2A^2 + m^2c^4. \tag{13}$$

Here, $\mathbf{A}$ represents the zeropoint electromagnetic field vector potential. In the rest frame of the particle, i.e., when the velocity $\mathbf{v} = 0$, the above expression reduces to

$$E_0 \sim mc^2 + \frac{e^2A^2}{2mc^2}, \tag{14}$$

where the higher order terms are neglected. Thus, under the influence of zeropoint field, the term $e^2A^2/2mc^2$ in the above equation gives a correction to mass. Expanding the vector potential in terms of its creation and annihilation operators and averaging in the standard form, it can be shown that the correction term [23]

$$\frac{e^2}{2mc^2}\langle A^2\rangle = \frac{\alpha}{2\pi}\frac{(\hbar\omega_c)^2}{mc^2}, \tag{15}$$

where ω_c is the cut-off frequency and α is the fine structure constant. When the cut-off frequency is equal to the frequency of oscillation of the particle, $\omega_c = \omega_0$ and using Einstein-de Broglie formula $\hbar\omega_0 = mc^2$, the mass correction can be expressed in the following form

$$\langle\Delta E_0\rangle = \delta mc^2 = \frac{\alpha}{2\pi}\,mc^2. \tag{16}$$

Thus in the presence of zeropoint field, the vector potential term in (15) gives the mass correction and it was obtained by Schwinger in quantum electrodynamics. The particle mass which arises due to local complex rotations in the zeropoint field is regarded as the so called bare mass and when the particle is observed from an arbitrary frame, the particle mass has some mass correction due to the presence of external zeropoint field.

3 Equation of motion of the particle with complex structure

It should be noted that, (1) contains the so called runaway and causal problems. In the Landau approximation, the damping term is written as a derivative of external force. In this case, the runaway and causal problems are eliminated and the exact equation of motion of a charged particle was recently given by Rohlrich [34] and Yaghjian [35]. In the equation of motion of the charged particle, centre of mass appears as if the total charge is at that point. In other words, there is no distinction between centre of mass and centre of charge points. In the case of extended particle structure, it has been clarified in the previous sections that the external zeropoint field must be responsible for the internal complex rotations and at the same time for the deviations in the path of the particle when it is moving with certain velocity. The external zeropoint field is then expressed as a function of complex vector X, $\mathbf{E}_z = \mathbf{E}_z(X,t)$ and expanding it gives

$$\mathbf{E}_z(X,t) = \mathbf{E}_z(\mathbf{x},t) + \mathbf{i}\,\xi \left|\frac{\partial \mathbf{E}_z(\mathbf{x},t)}{\partial \mathbf{x}}\right|_{\mathbf{x}\to 0} + O(\xi^2). \tag{17}$$

The second term on right hand side of the above equation is independent of x and it is a function of ξ only. Neglecting higher order terms in (17) and representing the second term on right by $\mathbf{i}\,\mathbf{E}_z(\xi,t)$, the external zeropoint field $\mathbf{E}_z(X,t)$ can be decomposed into a vector and a bivector parts

$$\mathbf{E}_z(X,t) = \mathbf{E}_z(\mathbf{x},t) + \mathbf{i}\,\mathbf{E}_z(\xi,t). \tag{18}$$

The random fluctuations produce kicks in all directions and leads to random fluctuations of the centre of mass point and at the same time random fluctuations also produce internal complex oscillations or rotations. Thus the force acting on the charged particle can be decomposed into two terms, the force acting on the centre of mass and the force acting on the centre of charge. For the field acting on the centre of mass, the particle mass and charge appear as if they are at the centre of mass point and we treat the equation of motion of the particle in the point particle limit. However, for the field acting on the centre of charge, the effective mass seen by the zeropoint field is the mass due to the potential $U_z \sim e^2/2R \sim m_z c^2$. The magnitude of R is of the order of Compton wavelength. Then the effective mass m_z in the zeropoint field is approximately equal to the electromagnetic mass which is proportional to the electromagnetic potential due to charge e at the center of mass position. Replacing the position vector $\mathbf{x}$ by the complex vector X and $\mathbf{E}_z(x,t)$ by the complex field vector $\mathbf{E}_z(X,t)$ in (1) and separating vector and bivector parts gives the equations of motion of the centre of mass and the centre of charge respectively. The equation of motion of center of mass is the Abraham-Lorentz equation of motion of a charged point particle in the external electromagnetic zeropoint field given by (1) and the motion of the centre of mass of the particle is observed from an arbitrary frame of reference. In the rest frame of the particle, the equation of motion represents the equation of motion of center of charge

$$m_z\ddot{\xi} - \Gamma_z m_z \ddot{\mathbf{u}} + m_z\omega_0^2\xi = e\mathbf{E}_z(\xi,t). \tag{19}$$

The terms $\Gamma_z m_z\ddot{\mathbf{u}}$ and $m_z\omega_0^2\xi$ are radiation damping and binding terms respectively. The damping constant in the above equation is defined as $\Gamma_z = (2e^2)/(3m_z c^3)$.

4 Average zeropoint energy associated with the particle in its rest frame

The zeropoint field and particle interaction takes place at resonance and the particle oscillates at resonant frequency ω_0. In other words, the particle oscillator absorbs energy from the zeropoint field at a single frequency which is the characteristic frequency of oscillation. Since, both radiation damping and binding terms are much smaller than the force term in (19) one can neglect these terms and integrating with respect to time t gives the internal velocity of rotation of the particle

$$\mathbf{u}(t) = \frac{e}{m_z}\int_0^\tau \mathbf{E}_z(\xi,t)dt. \tag{20}$$

Here, the upper limit of integration is chosen as the characteristic time τ required by the electromagnetic wave to traverse a distance equal to the size of the particle. The electric field vector $\mathbf{E}_z(\xi,t)$ is expressed in the same form as that of Rueda [19],

$$\mathbf{E}_z(\xi,t) = \sum_{\lambda=1}^{2}\int d^3k\,\epsilon(\mathbf{k},\lambda)\frac{H(\omega)}{2}\times \left[a e^{i(\mathbf{k}.\xi-\omega t)} + a^* e^{i(\mathbf{k}.\xi-\omega t)}\right], \tag{21}$$

where $a = \exp(-i\theta(\mathbf{k},\lambda))$, $a^* = \exp(i\theta(\mathbf{k},\lambda))$ and $\epsilon(\mathbf{k},\lambda)$ is the polarization vector and the normalization constant is set equal to unity. The phase angle $\theta(\mathbf{k},\lambda)$ is a set of random variables uniformly distributed between 0 and 2π and are mutually independent for each choice of wave vector $\mathbf{k}$ and λ. The stochastic nature of the field lies in these phase angles and a statistical average of these phase angles gives an effective value of the field. For point particles, because the size is zero, we find the spectral divergence of zeropoint field. However, for particles with extended structure, one can discern a natural cut-off wavelength associated with the particle size. The convergence factor gives an upper bound to the energy available

from the electromagnetic zeropoint field and it is associated with the characteristic function $H(\omega)$ of the zeropoint field. The function $H(\omega)$ is given by $2\pi^2 H^2(\omega) = \eta(\omega)\hbar\omega$. In (21), integrating the electric field vector with respect to time gives

$$I = \sum_{\lambda=1}^{2} \int_0^\infty d^3k \, \frac{H(\omega)}{2} \left[\epsilon(\mathbf{k}, \lambda)\, a e^{i\mathbf{k}.\xi} \left(\frac{e^{-i\,\omega\tau} - 1}{-i\,\omega} \right) + \epsilon(\mathbf{k}, \lambda)\, a e^{i\mathbf{k}.\xi} \left(\frac{e^{-i\,\omega\tau} - 1}{-i\,\omega} \right) \right].$$

The charge current in the rest frame of the particle is the charge times the internal velocity of the particle. The interaction energy of the charged particle with the zeropoint field is expressed as the charge current times the vector potential of the zeropoint field. However, one can express the vector potential as the integral of the zeropoint electric field vector. Then the average zeropoint energy acquired by the particle is expressed as

$$\langle E_0 \rangle = \frac{e^2}{m} \langle II^* \rangle. \tag{22}$$

The averages of random phase and the polarization vector are expressed as follows

$$\langle aa^* \rangle = \delta(\lambda - \lambda')\, \delta^3(k - k')\,;\ \langle aa \rangle = 0\,;\ \langle a^* a^* \rangle = 0$$
$$\langle \epsilon(\mathbf{k}, \lambda) \epsilon^*(\mathbf{k}, \lambda) \rangle = \delta_{ij} - \tfrac{k_1}{k^2}$$
$$\textstyle\sum_{\lambda=1}^{2} \int d^3k \, \langle \epsilon(\mathbf{k}, \lambda) \epsilon^*(\mathbf{k}, \lambda) \rangle = \tfrac{8\pi}{3} \int \omega^2 \, d\omega \,.$$

Using these stochastic averages, replacing the convergence factor by $\eta(\omega_0)$ and setting the upperlimit of integration to the frequency of oscillations in (22) gives

$$\langle E_0 \rangle = \frac{4e^2\hbar}{3\pi m_z c^3}\, \eta(\omega_0) \int_0^{\omega_0} \omega(1 - \cos\omega\tau)\, d\omega. \tag{23}$$

For an extended particle structure $\omega_0\tau = 2\pi$ and the above equation after integration reduces to

$$\langle E_0 \rangle = \eta(\omega_0)\, \frac{\Gamma_z \hbar \omega_0^2}{\pi}. \tag{24}$$

This result is similar to the result obtained by Reuda [19] and Puthoff [36]. However, the difference is that the damping constant is now replaced by Γ_z and cut-off frequency ω_c is replaced by the resonant internal frequency of oscillation of the particle. In (24), both the values for m_z and $\eta(\omega_0)$ are not known exactly and must be approximated. Instead, one can approximate $\eta(\omega_0)\Gamma_z\omega_0 \sim 1$ for the particle with extended structure. Then the average zeropoint energy acquired by the particle in its rest frame is

$$\langle E_0 \rangle = \frac{\hbar\omega_0}{\pi}. \tag{25}$$

This energy is similar to the zitterbewegung energy of Dirac electron in quantum mechanics.

5 Equation of motion of the particle with complex structure

In the above procedure, initially we have considered the charged particle without any mass. Such particle interacting with zeropoint field acquires mass due to particle resonant oscillations and gains energy from the electromagnetic zeropoint field. This average zeropoint energy of the particle appears as the mass of the particle. In the complex vector formalism of internal harmonic oscillator in zeropoint field, it has been shown by the author that the average energy $\langle E_0 \rangle$ is related to the mass through particle spin and represents the mass generated from the local complex rotations produced by the interaction of zeropoint field with the particle. The relation between average zeropoint energy and particle spin is given by the expression [26]

$$\langle E_0 \rangle - \omega_0 \langle \mathbf{s} \rangle = 0. \tag{26}$$

Let us denote $\omega_s = 2\omega_0$ and write the angular velocity bivector as $\Omega_s = -\mathbf{i}\,\sigma_s\omega_s$, where σ_s is a unit vector along the direction of spin. The average value of spin is obtained by taking the average over a half cycle, $\langle \mathbf{s} \rangle = \frac{2}{\pi}\mathbf{s}$. Substituting this average value of spin and $\langle E_0 \rangle$ from (25) in (26) gives the relation between particle mass and spin

$$mc^2 = \sigma_s \Omega_s . S, \tag{27}$$

where the relation $\hbar\omega_0 = mc^2$ is used and the bivector spin $S = \mathbf{i}\,\sigma_s\hbar/2$. The unit vector σ_s acting on an idempotent $\mathcal{J}_+ = (1 + \sigma_S)/2$ gives an eigenvalue +1. This statement is represented by an equation $\sigma_s\mathcal{J}_+ = +1\mathcal{J}_+$. When (27) is multiplied from right by an idempotent $\mathcal{J}_+$ on both sides the unit vector is absorbed by the idempotent and equating the scalar parts gives

$$mc^2 = \Omega_s . S. \tag{28}$$

Thus the mass of the particle turns out to be the local internal rotational energy given by the term $\Omega_s . S$. Since, the magnitude of spin and velocity of light are constants, the value of particle mass depends on the frequency of spin rotation and the different particles may have different frequencies of spin rotation. The above analysis shows that the internal complex rotation is responsible for the existence of particle mass. Then, one may initially consider a massless charged particle and it may acquire mass from zeropoint field through a local complex rotation.

When the particle is observed from an arbitrary frame of reference, the center of mass point moves with velocity $\mathbf{v}$. The equation of motion of centre of mass point is given by (1) and solving it by assuming the radiation damping and binding terms as small when compared to the force term, one can obtain the zeropoint energy absorbed by the point particle and it is given by (3). The cut-off frequency ω_c is the limiting frequency in the integration. When we assume the cut-off

frequency $\omega_c = \omega_0$ [37, 38] and after introducing the convergence factor $\eta(\omega_0) \sim 3/4$ in (3), the average energy representing the mass correction of the particle in the zeropoint field can be expressed as

$$\delta m = \frac{\alpha}{2\pi} mc^2. \quad (29)$$

This mass correction is too small and found to be similar to the expression found in quantum electrodynamics to the first order in the fine structure constant α.

6 Conclusions

In the stochastic electrodynamics with spin, it has been shown that the average zeropoint energy absorbed by the particle due to its internal motion gives the particle mass. When the particle center of mass point moves with certain velocity, we find the average energy absorbed by the particle gives the mass correction. In deriving both particle mass and mass correction, a convergence factor has been introduced for an extended particle. To understand the mechanism of mass generation of an elementary particle, one may initially assume a massless charged particle with complex structure and such a particle can be visualized as an oscillator in the fluctuating zeropoint field. Then the average energy absorbed by the oscillator refers to the particle mass. Finally, we conceive the idea that an elementary particle acquires mass from the interaction of ubiquitous zeropoint field.

Submitted on March 18, 2016 / Accepted on March 21, 2016

References

1. Weyssenhoff J., Raabbe A. Relativistic dynamics of spin fluids and spin particles. *Acta. Phys. Pol.*, 1947, v. 9, 7.
2. Barut A. O., Zanghi A. J. Classical model of the Dirac electron. *Phys. Rev. Lett.*, 1984, v. 52, 2009–2012.
3. Salesi G., Recami E. A velocity field and operator for spinning particles in (nonrelativistic) quantum mechanics. *Found. Phys.*, 1998, v. 28, 763–773.
4. Pavsic M., Recami E., Rodrigues W. A., Maccarrone G. D., Raciti F., Saleci G. Spin and electron structure. *Phys. Lett. B.*, 1993, v. 318, 481.
5. van Holten J. W. On the electrodynamics of spinning particles. *Nuclear Phys. B.*, 1991, v. 356, 3–26.
6. Hestenes D. Zitterbewegung in quantum mechanics. *Found. Phys.*, 2010, v. 40, 1–54.
7. Muralidhar K. Classical origin of quantum spin. *Apeiron*, 2011, v. 18, 146.
8. Muralidhar K. The spin bivector and zeropoint energy in geometric algebra. *Adv. Studies Theor. Phys.*, 2012, v. 6, 675–686.
9. Haisch B., Rueda A., Dobyns Y. Inertial mass and quantum vacuum fields. *Ann der Physik.*, 2001, v. 10, 393–414.
10. Sidharth B. G. Revisiting zitterbewegung. *Intl. J. Theor. Phys.*, 2009, v. 48, 497–506.
11. Milonni P. W. The Quantum Vacuum: An Introduction to Quantum Electrodynamics. Academic Press, Boston, 1994.
12. Boyer T. H. The classical Vacuum. *Sci. Am.*, 1985, v. 253, 70.
13. de La Pena L., Cetto A. M. The Quantum Dice – An Introduction to Stochastic Electrodynamics. Kluwer Academic Publishers, Dordrecht, 1996.
14. Marshall T. W. Random electrodynamics. *Proc. Roy. Soc. A.*, 1963, v. 276, 475–491.
15. Boyer T. H. Random electrodynamics – The theory of classical electrodynamics with classical electromagnetic zero point radiation. *Phys. Rev. D.*, 1975, v. 11, 790.
16. Boyer T. H. Unretarded London-van der Waals forces derived from classical electrodynamics with classical electromagnetic zeropoint radiation. *Phys. Rev.*, 1972, v. 6, 314.
17. Boyer T. H. Quantum zeropoint energy and long range forces. *Annals of Phys.*, 1970, v. 56, 474.
18. Boyer T. H. Connection between the adiabatic hypothesis of old quantum theory and classical electrodynamics with classical electromagnetic zero-point radiation. *Phys. Rev. A.*, 1978, v. 18, 1238.
19. Rueda A. Behaviour of classical particles immersed in electromagnetic zero-pint field. *Phys. Rev. A.*, 1981, v. 23, 2020.
20. Cavalleri G., Barbero F., Bertazzi G., Cesaroni E., Tonni E., Bosi L., Spavieri G., Gillies G. T. A qualitative assessment of stochastic electrodynamics with spin (SEDS): Physical principles and novel applications. *Front. Phys. China.*, 2010, v. 5, 107–122.
21. de la Pena L., Cetto A. M., Hernandez A. V. The Emerging Quantum: The Physics Behind Quantum Mechanics. Springer, Cham, 2015.
22. Lorentz H. A. The Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat. G. E. Stechert and Co., New York, 1916.
23. Haitch B., Reuda A., Puthoff H. E. Inertia as a zero point field Lorentz force. *Phys. Rev. A.*, 1994, v. 49, 678–894.
24. Cavalleri G. $\hbar$ derived from cosmology and origin of special relativity and QM. *Nuovo Cimento. B.*, 1997, v. 112, 1193–1205.
25. Hestenes D. Oersted Medal Lecture 2002: Reforming the Mathematical Language of Physics. *Am. J. Phys.*, 2003, v. 71, 104–121.
26. Muralidhar K. Complex vector formalism of harmonic oscillator in geometric algebra: Particle mass, spin and dynamics in complex vector space. *Found. Phys.*, 2014, v. 44, 265–295.
27. Jammer M. Concepts of Mass in Contemporary Physics and Philosophy. Princeton University Press, Princeton, New Jersey, 2000.
28. Wentzel G. Quantum Theory of Fields. Dover Publications Inc., Mineola, New York, 2003.
29. Modanese G. Inertial mass and vacuum fluctuations in quantum field theory. *Found. Phys. Lett.*, 2003, v. 16, 135–141.
30. Pollock M. D. On vacuum fluctuations and particle masses. *Found. Phys.*, 2012, v. 42, 1300–1338.
31. Ibison M. Massless classical electrodynamics. *Fizika. A.*, 2003, v. 12, 55–74.
32. Ibison M. A Massless Classical Electron. In: Chubykalo A., Espinoza A., Smirnov-Rueda R., Onoochin V., eds. Has the Last Word Been Said on Classical Electrodynamics? – New Horizons. Rinton Press, New Jersey, 2003.
33. Doran C., Lasenby A. Geometric Algebra for Physicists. Cambridge University Press, Cambridge, 2003.
34. Rohrlich F. Classical Charged Particles. World Scientific, Singapore, 2007.
35. Yaghjian A. D. Relativistic Dynamics of a Charged Sphere. Springer, New York, 2006.
36. Puthoff H. C. Gravity as zeropoint fluctuation force. *Phys. Rev. A.*, 1989, v. 39, 2333–2342.
37. Masperi L. Qualitative description of electron's anomalous magnetic moment. *Revista Brasillera de Fisica.*, 1989, v. 19, 215.
38. Bjorken J. D., Drell S. D. Relativistic Quantum Mechanics. McGraw-Hill, New York, 1964.

The Relationship Between the Possibility of a Hidden Variable in Time and the Uncertainty Principle

Eyal Brodet
E-mail: eyalbrodet@hotmail.com

In this paper we will discuss the relationship between the possibility of a hidden variable in time and the uncertainty principle. The discussion consists in a fundamental look at the decay time processes of unstable elementary particles. As will be argued, the hidden variable in time possibility may result in a possible way to bypass the energy-time uncertainty principle. Therefore energy and time information may be known simultaneously in the decay time process. A fundamental and general experimental way to test the above is suggested.

1 Introduction

In this paper we investigate the connection between the possibility of a hidden variable in time and the uncertainty principle. In [1] the process of $e^+e^- \to \mu^+\mu^-$ was discussed and it was questioned how come under what appears to be identical local initial conditions we get a distribution of decay time values for the μ^+ and the μ^-. In[1] there was no discussion about the muon mass width that in fact means that the muons are in principle not completely identical to each other and therefore the local initial conditions are not completely identical between the different events. In [1] it was assumed that the muon mass width could not explain (at least not by itself) the exponential decay time distribution of the muons. Therefore the suggestion was that there exists another internal property within the muons that is responsible for generating the muon decay time distribution.

However this could not be the complete explanation as the muons do have a narrow distribution of mass values and therefore there are slightly different local initial conditions in this process between different events. This fact has to be taken into account in a complete explanation for the decay time distribution in this process. As the muon mass width is part of the uncertainty principle, in this paper we will discuss the connection of the uncertainty principle to the hidden variable in time possibility and attempt to incorporate the two. Moreover we will discuss how the existence of a hidden variable in time could help to bypass the energy-time aspect of the uncertainty principle. Finally an experimental way to test the above is discussed.

This paper is organized as follows. Section 2 discusses the theoretical background. Section 3 describes a possible experimental way to bypass the energy-time uncertainty principle in case a hidden variable in time do exist. The conclusions are presented in section 4.

2 Theoretical discussion

2.1 Background

In this paper we investigate the connection between the possibility of a hidden variable in time and the uncertainty principle. The hidden variable in time possibility first presented in [1] gives a deterministic approach that attempts to explain the distribution in decay time as a result of a compatible distribution in an additional internal property within the particles. The suggestion was that this additional internal property is related to the frequency of the virtual boson emission and absorption and therefore as it is related to time and affects the decay time of particles it was termed a hidden variable in time (f_r). However even if the above is correct this could not be a complete explanation as we have to take into account a known distribution in the initial decaying particles which is the distribution in their mass.

In [1] it was assumed that this mass distribution of for example the muon particles can not solely explain the muon exponential decay time distribution. More specifically it was assumed that the Breit-Wigner distribution of the mass value could not be translated in a deterministic, unique and logical way, using the Standard Model, into the exponential decay time distribution that we observe. One could convince oneself intuitively that this is the case by considering the peaks of the two distributions which are at $m = M_{mean}$ for the Breit-Wigner case and $t = 0$ for the exponential decay case and also the tails which are for the mass Breit-Wigner case at $m = 0$, and at $m = \infty$ (two tails) and for decay time case at $t = \infty$.

Therefore one can not get a logical connection between the two distributions because one could not associate the process initial and final condition logically considering what we know from the Standard Model. That is if we start from the two peaks as the most common and popular initial conditions where most of the events are then we get two different initial conditions for the mass value at the tails ($m = 0,\ m = \infty$) that give a single final condition which is the tail of the exponential at $t = \infty$.

This does not give a logical and deterministic explanation as logically under different initial conditions we should get different final conditions considering the dependence of the decay time on the mass as described in the Standard Model and the experimental decay time results, i.e the higher the mass is the shorter the decay time is. Therefore the mass distribution could not generate deterministically and logically

using the Standard Model, the observed muon decay time distribution, and we need an alternative explanation. Perhaps in the form of the hidden variable f_r.

The standard model does however, give a general link between the Breit-Wigner shape and the exponential decay time shape for a given particle which narrows down the uncertainty by telling us the favorite mass value of the particles is M_{mean} and that an enhanced fraction of them will decay almost instantly after they are born. For example if 40% of the masses are at a bin around M_{mean} then $40 * 40 = 16\%$ of them will decay in the first decay time bin in the decay time distribution.

That is if we measure a specific mass to be on the mean value then we know that there is a 16% probability that the particle would decay in the first decay time bin. This is compared to what we know from the uncertainty principle, where knowing exactly the mass value yields a complete uncertainty on the decay time. Therefore the Standard Model reduces the uncertainty with respect to the uncertainty principle by allowing us to calculate the Breit-Wigner and the exponential decay time distributions.

2.2 $\Delta m, \Delta t$ and f_r

The distribution of Δm is known from the Breit-Wigner but the distribution of Δt is experimentally unknown (we do not know how to deduce it from the exponential decay distribution), we only know the maximum value of it from the knowledge of Δm and the boundaries given by the uncertainty principle.

According to [1], if one knows the true particle decay $T1$, time then one may know f_{r1} from Fig. 1. In this case this particular f_{r1} has two possible mass value $M1, M2$ as shown in Fig. 2. These two mass values may have the same value of f_{r1} but with very different mean lifetimes given for example in the muon case, from the known Standard Model formula:

$$\tau_{\mu(1,2)} = \frac{192\,\pi^3 h^7}{G_f M(1,2)_\mu^5\, c^4}\,. \qquad (1)$$

where G_f is the Fermi coupling constant. In the case when f_r exists, one may have a deterministic link between $\Delta m, \Delta t, f_r$ and t which may cancel the uncertainty limitations as will be discussed later on. Without f_r there is no deterministic link between a specific mass $M1$ and a specific decay time $T1$.

If f_r exists the exponential shape is the slope of the f_r depending on the mean lifetime which gives a deterministic description for a specific event using extra information in the form of f_r.

2.3 Mathematical relationship between f_r and the exponential and Breit-Wigner distributions

Putting the above into a mathematical form gives us two expressions for f_r:

$$f_{r(i)} = f(m_i)\,A_i = f(m_i)\,\exp\left(-\frac{t_i}{\tau_i}\right) \qquad (2)$$

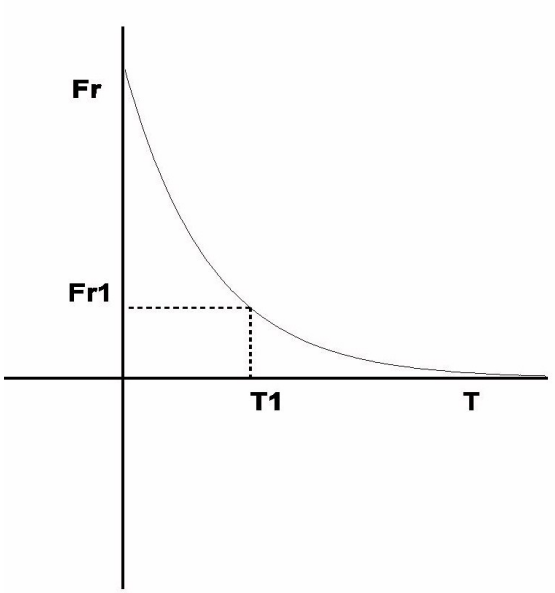

Fig. 1: f_r versus the distribution of the particles decay time.

$$f_{r(i)} = f(m_i)\,f(E_i) = f(m_i)\,\frac{k_1}{(E_i^2 - M_{mean}^2)^2 + M_{mean}^2\,\Gamma_i^2} \qquad (3)$$

where $f(m_i)$ is the mass amplitude and $E_i = m_i\,c^2$. One possibility for $f(m_i)$ may be: $f(m_i) = M_{mean} + (m_i - M_{mean})\,k_2$ where k_1 and k_2 are parameters with yet unknown values. The above relationship suggests the following:

For mass measured close to the muon mean mass values Δm is small and Δt is large (but we cannot be sure at this stage how large as we do not know how to deduce the uncertainty Δt from the exponential muon decay time distribution). We can only get the maximum value for Δt for small Δm using the uncertainty principle. From Fig. 2 we can see that small Δm values correspond to high f_r values and therefore, as can be seen from Fig. 1 to short decay time values.

For mass measured far from the mean muon mass value (lower or bigger), we know that Δm is large and Δt is small. Again we cannot be sure for a particular event how small is Δt, however we only know that it has to be small as Δm is large in order to satisfy the uncertainty principle. From Fig. 2 we can see that large Δm corresponds to low f_r values and therefore, as can be seen from Fig. 1, to long decay time values.

Therefore the effect of the uncertainty principle assuming the existence of f_r, on the decay process is that it associates a particular and different uncertainty on each decay time and mass values. This is where masses around M_{mean} are assumed to have short decay times and have small Δm and large Δt, and masses that are away from M_{mean} (smaller or greater) are assumed to have larger decay times and larger Δm and smaller Δt.

This is where the limitation associated by the uncertainty principle of knowing simultaneously the exact mass and de-

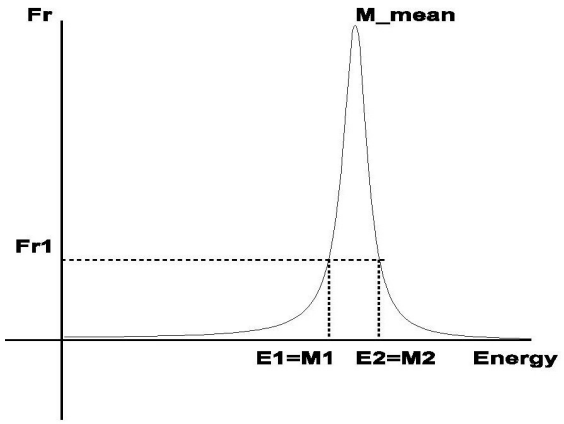

Fig. 2: f_r versus the distributions in the particles mass.

cay time of a particle still remains. In the next section we will discuss how this limitation could be bypassed.

3 Possible bypass of the energy-time uncertainty principle

The possibility of a hidden variable in time opens up a new way to fundamentally bypass the above limitation on the simultaneous knowledge of the muon mass and decay time values. This could be expressed by the following measurement that may be done by known detectors [2]:

As measuring the muon mass exactly is experimentally difficult due to the missing energy of the neutrino involved, we may turn to measuring exactly its decay time.

Therefore, if we measure a specific muon decay time $t1$, we know its f_r from Fig. 1. Therefore we could know its two associated masses $m1$ and $m2$ from Fig. 2 and its mass uncertainty Δm. From the uncertainty principle we could then also know its maximum Δt uncertainty. Therefore this gives us three possible decay time $t1$, $t1 - \Delta t$, $t1 + \Delta t$ and six possible masses that are associated to these three decay times. Now we need to decide which pair of decay time and mass values is the correct one for that particular event. We can attempt to do that by measuring exactly the muon electric charge Q in that event from:

$$\frac{M V^2}{r} = QVB \qquad (4)$$

where B is the external magnetic field, M is the muon mass and we can measure the momentum from the curvature r and the velocity V from the Cherenkov detector. Therefore after knowing the charge we may deduce the factor $A = \exp(-t/\tau)$ according to [3]. This A value corresponds to a part of the particles f_r in that particular event where $f_{r(i)} = f(m_i) A_i$ as shown in (2). Now all we need to do is to see which pair of mass and decay time values is closest in value to the measurement of the factor A and find the correct initial mass and final decay time in that particular event, thereby bypassing the uncertainty principle.

4 Conclusion

The implication of a hidden variable in time on the energy-time uncertainty principle was discussed. A fundamental way was presented to bypass the uncertainty principle through measuring the decay time and charge value in a specific $e^+e^- \rightarrow \mu^+\mu^-$ event, thereby knowing the exact value of the initial muon mass and decay time.

Submitted on April 2, 2016 / Accepted on April 6, 2016

References

1. Brodet E. The possibility of a hidden variable in time. *Physics Essays*, 2010, v. 23 (4), 613–617.
2. The DELPHI detector at LEP. http://delphiwww.cern.ch/. Accessed Dec 2015.
3. Brodet E. Incorporating the possibility of a hidden variable in time in the Standard Model and the experimental implications. *Physics Essays*, 2016, v. 29 (1), 57–61.

A Modern Interpretation of the Dirac-Electron Continuity Equation

William C. Daywitt

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

This paper re-derives the Dirac continuity equation for the electron from the viewpoint of the Planck vacuum (PV) theory. Results show the equation to be a spacetime equation (whose line elements are cdt and dx^k) that equates the normalized ct-gradient of the probability density ($\psi^\dagger\psi$) to the normalized negative divergence of the quantity ($\psi^\dagger\boldsymbol{\alpha}\psi$).

1 Introduction

The Dirac equation that defines the free-electron spinor field $\psi = \psi(\mathbf{r}, t)$ [1, p.74]

$$ic\hbar\frac{\partial\psi}{c\partial t} = (c\boldsymbol{\alpha}\cdot\widehat{\mathbf{p}} + mc^2\beta)\psi \tag{1}$$

where $\widehat{\mathbf{p}}$ ($=-i\hbar\nabla$) is the vector momentum operator, can be expressed as

$$ic\hbar\left(\frac{\partial}{c\partial t} + \boldsymbol{\alpha}\cdot\nabla\right)\psi = mc^2\beta\psi \tag{2}$$

where c is the speed of light, $\hbar$ is the reduced Planck constant, and m is the electron mass. The spinor field ψ is the 4×1 column vector

$$\psi = \begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix}. \tag{3}$$

The two 4x4 matrices in (1) and (2) are defined by

$$\alpha_k = \begin{pmatrix}0 & \sigma_k\\ \sigma_k & 0\end{pmatrix} \quad\text{and}\quad \beta = \begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix} \tag{4}$$

where $k = (1, 2, 3)$ and I is the 2×2 unit matrix. The three 2×2 Pauli spin matrices σ_k are [1, p. 12]

$$\sigma_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad \sigma_2 = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\quad \sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \tag{5}$$

and the operator on the left side of (2) reduces to

$$\left(\frac{\partial}{c\partial t} + \boldsymbol{\alpha}\cdot\nabla\right) = \left(\frac{\partial}{c\partial t} + \sum_{k=1}^{3}\alpha_k\frac{\partial}{\partial x^k}\right). \tag{6}$$

In its rest frame the massive electron core $(-e_*, m)$, with its zero-point derived mass m [2], exerts the two-term coupling force [3, Sec. 7-8]

$$F(r) = \frac{e_*^2}{r^2} - \frac{mc^2}{r} = \frac{(-e_*)(-e_*)}{r^2} - \frac{mm_*G}{r_*r} \tag{7}$$

on the PV quasi-continuum, where e_* is the massless bare charge and $G\,(=\,e_*^2/m_*^2)$ is Newton's gravitational constant. The first $(-e_*)$ in (7) belongs to the electron and the second to the separate Planck particles making up the degenerate PV state. The two terms in (7) represent respectively the Coulomb repulsion between the electron charge and the separate PV charges, and the second their mutual gravitational attraction.

The particle/PV coupling force (7) vanishes at the electron Compton radius $r_c\,(=\,e_*^2/mc^2)$. In addition, the vanishing of $F(r_c)$ is a Lorentz invariant constant [4] that leads to the important Compton-(de Broglie) relations

$$r_c\cdot mc^2 = r_d\cdot cp = r_L\cdot E = r_*\cdot m_*c^2 = e_*^2 \quad (=c\hbar) \tag{8}$$

where $r_d = r_c/\beta_0\gamma_0$ and $r_L = r_c/\gamma_0$, and $r_*\,(=\,e_*^2/m_*c^2)$ and m_* are the Compton radius and mass of the Planck particles within the PV state. The ratio of the electron speed v to the speed of light c is β_0 and $\gamma_0 = 1/(1-\beta_0^2)^{1/2}$. The relativistic momentum and energy following from the invariance of $F(r_c) = 0$ are $p\,(=\,m\gamma_0 v)$ and $E\,(=\,m\gamma_0c^2)$, from which $E = (m^2c^4 + c^2p^2)^{1/2}$ is the relativistically important energy-momentum relationship.

Using (8), (2) can be expressed as

$$ie_*^2\left(\frac{\partial}{c\partial t} + \boldsymbol{\alpha}\cdot\nabla\right)\psi = mc^2\beta\psi \tag{9}$$

or

$$ir_c\left(\frac{\partial}{c\partial t} + \boldsymbol{\alpha}\cdot\nabla\right)\psi = \beta\psi \tag{10}$$

where the partial derivatives within the parentheses are normalized by the Compton radius r_c. The spinor field that is the hermitian conjugate of ψ is the 1×4 row vector $\psi^\dagger = (\psi_1^\dagger, \psi_2^\dagger, \psi_3^\dagger, \psi_4^\dagger)$. Then, pre-multiplying (10) by $\psi^\dagger$ leads to

$$ir_c\psi^\dagger\left(\frac{\partial}{c\partial t} + \boldsymbol{\alpha}\cdot\nabla\right)\psi = \psi^\dagger\beta\psi\,. \tag{11}$$

Taking the hermitian conjugate of (10), post-multiplying by ψ, then yields [1, p. 76]

$$-ir_c\left(\frac{\partial}{c\partial t} + \boldsymbol{\alpha}\cdot\nabla\right)\psi^\dagger\psi = \psi^\dagger\beta\psi\,. \tag{12}$$

Subtracting (12) from (11) finally leads to the continuity equations [1, p. 76]

$$ir_c\left[\frac{\partial(\psi^\dagger\psi)}{c\partial t} + \nabla\cdot(\psi^\dagger\boldsymbol{\alpha}\psi)\right] = 0 \tag{13}$$

or

$$\frac{\partial(\psi^\dagger\psi)}{c\partial t/r_c} + \sum_{k=1}^{3} \frac{\partial(\psi^\dagger\alpha_k\psi)}{\partial x^k/r_c} = 0 \qquad (14)$$

for the electron. From (8), the presence of r_c in these two equations connects the electron core dynamics to a wave traveling within the vacuum state [5].

2 Comments and Conclusions

Dividing (13) by ir_c yields the equation

$$\frac{\partial(\psi^\dagger\psi)}{\partial t} + \nabla\cdot(\psi^\dagger c\boldsymbol{\alpha}\psi) = 0 \qquad (15)$$

where the 4×4 matrix $c\boldsymbol{\alpha}$ looks like a velocity operator because of the speed of light c. This observation then leads intuitively to the standard continuity equation [1, p. 76]

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0 \qquad (16)$$

where $\rho = \psi^\dagger\psi$ is the probability density and $j^k = \psi^\dagger c\alpha_k\psi$ is the kth component of the probability current density. Integrating (16) over the volume V (assumed to contain the electron core $(-e_*, m)$), and using the divergence theorem, leads to [1, p. 77]

$$\frac{\partial}{\partial t}\int_V d\rho\, d^3x + \int_S \mathbf{j}\cdot d\vec{S} = 0\,. \qquad (17)$$

where the surface S surrounds the volume V.

So far, so good. But there is a problem: treating $c\boldsymbol{\alpha}$ as a free-space matrix velocity leads to a tortured interpretation of that operator that cries out for a better explanation. From the PV perspective, that explanation is apparent from equation (14)

$$\frac{\partial(\psi^\dagger\psi)}{c\partial t/r_c} + \sum_{k=1}^{3} \frac{\partial(\psi^\dagger\alpha_k\psi)}{\partial x^k/r_c} = 0$$

where the Minkowski-like line elements, cdt and dx^k associated with the partial derivatives, are normalized by the electron Compton radius r_c. The form of this equation suggests that it is associated with a distorted spacetime [6, p. 27] (the distortion coming from the r_c and the α_k), rather than a free-space velocity dynamic. Furthermore, the absence of the dynamical electron parameters p and E from (8), and the fact that $c\boldsymbol{\alpha}$ is not a recognizable free-space operator, suggest that (14) refers to a PV substructure dynamic [7] (driven by the electron core dynamic), where the normalized ct-gradient of $(\psi^\dagger\psi)$ equals the normalized negative divergence of $(\psi^\dagger\boldsymbol{\alpha}\psi)$.

Finally, the assumption that the PV is a degenerate state implies that the Planck-particle energy eigenstates are full. So if there is a current wave propagating within the PV, it cannot involve a Planck particle current (because the Planck particles are not free to move macroscopically). Thus $c\boldsymbol{\alpha}$ must refer, in part, to a localized percussion-like spinor wave within that vacuum state, analogous to a wave traveling on the surface of a kettle drum.

Equations (13) and (14) and the previous two paragraphs represent the PV view of the Dirac-electron continuity equation.

Submitted on April 15, 2016 / Accepted on April 30, 2016

References

1. Gingrich D.M. Practical Quantum Electrodynamics. CRC, The Taylor & Francis Group, Boca Raton, London, New York, 2006.
2. Daywitt W.C. Why the Proton is Smaller and Heavier than the Electron. *Progress in Physics*, 2014, v. 10, 175.
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, no. 1-1, 22.
4. Daywitt W.C. The de Broglie Relations Derived from the Electron and Proton Coupling to the Planck Vacuum State. *Progress in Physics*, 2015, v. 11(2), 189.
5. Daywitt W.C. The Dirac-Electron Vacuum Wave. *Progress in Physics*, 2016, v. 12(3), 222.
6. Leighton R.B. Principles of Modern Physics. McGraw-Hill Book Co., New York, Toranto, London, 1959.
7. Daywitt W.C. The Strong and Weak Forces and their Relationship to the Dirac Particles and the Vacuum State. *Progress in Physics*, 2015, v. 11(1), 18.

Optics of the Event Horizon Telescope

Trevor W. Marshall
Buckingham Centre for Astrobiology, The University of Buckingham, Buckingham MK18 1EG, UK
E-mail: trevnat@talktalk.net

It is suggested in this article that part of the signal in the 1.3 mm range from Sagittarius A* originates inside the central collapsar, rather than coming entirely from its accretion disc. The suggestion has its origin in the discovery that the classic article of Oppenheimer and Snyder contains a basic error in its assertion that the light, from a collapsing object lying entirely within its own photonsphere, is progressively cut off as the object shrinks towards its gravitational radius, where a large part of the Oppenheimer-Snyder collapsar's material is concentrated. The signal from the collapsar has certain features which may make it possible to distinguish its image from that of the accretion disc.

1 Introduction

At the centre of our galaxy, 8 kp distant from us, there is an object named Sagittarius A* whose mass is 4.1 megasuns. It is popularly classified as a black hole, with a spherical* region of radius 1.2×10^7 km around it bounded by an "event horizon"; according to black-hole theory no light from Sagittarius A* can cross this horizon.

In two recent articles [1, 2] it was shown that there is a solution of the field equations of General Relativity for such a supermassive object, which has no singularity at $r = 0$, and which allows light signals to cross the horizon. The latter property of the solution was demonstrated for the case of rays which are normal to the event horizon, and the present article demonstrates that it may be extended to all orientations. In addition we consider the range of angles for which light originating at the surface of such a collapsar crosses the photonsphere, at 1.5 times the gravitational radius, and consequently may reach a terrestrial telescope. There is currently a project called the Event-Horizon Telescope [3] (EHT) designed to look at the signal from the neighbourhood of Sagittarius A* in the 1.3 mm range.

Central to the widespread belief in the validity of black-hole theory is the article of Oppenheimer and Snyder (OS) [4]. This reported, without giving details, an investigation of the light signal from a supermassive object, arriving at the following conclusion

> All energy from the surface of the star will be reduced very much in escaping ... by the gravitational deflection of light which will prevent the escape of radiation except through a cone about the outward normal of progressively shrinking aperture as the star contracts. The star thus tends to close itself off from any communication with a distant observer.

The property of the OS metric claimed by Penrose, which he needed as a prerequisite for his singularity theorem [5], was the stronger one known as the *trapped surface*. The publications cited above show that neither of these properties in fact holds for the OS metric.

In the following two sections we shall use precisely the OS metric to show that the progressively shrinking aperture of the emission cone has no effect on the size of the image of the collapsing object, and only a marginal effect on its total luminosity. This result leads us to suggest that the signal from Sagittarius A* comes partly from the surface of the collapsar itself, and not entirely from the accretion disc, as is assumed in most current analyses. The accretion disc may well have a substantially higher temperature than the collapsar, but that is probably offset by the vastly greater area of the latter. Note also that the millimetre range of wavelength investigated by the EHT corresponds to the maximum of a Planck spectrum of just a few degrees Kelvin; to support our analysis, the collapsar must retain only the merest relic of its thermal energy.

The OS article reached another conclusion, stated in their Abstract, namely

> ... an external observer sees the star asymptotically shrinking to its gravitational radius.

This result contradicts directly Penrose's description of the OS results and was verified by me in [1]. The point is that OS showed that there is a common system of coordinates applicable to both the exterior and interior of the collapsar. My article [1] demonstrated that the density distribution of the OS "dust cloud" becomes concentrated near the surface as it shrinks to the gravitational radius; no exotic process like the modern black-hole one of "spaghettification" [6] occurs when a notional spaceship crosses the event horizon. OS should be considered responsible for the notion that further shrinkage occurs within the gravitational radius only in so far as they gave their article the misleading title "On continued gravitational contraction".

It should be noted that in the exterior, and hence in what should now be recognized as the universal, time frame the collapsar's shrinkage to the gravitational radius takes an infinite lapse of time. We shall show in the following section that in the limit there is an underlying infinite red shift, which

*For the purposes of this article we ignore its spin.

causes not only the surface itself, but also all light signals approaching it, to be infinitely slowed down. This is the real significance of the event horizon, but it is my contention that a real collapsar, with an internal pressure resulting from the intervention of forces other than gravitational, stops shrinking before it reaches the gravitational radius. For example, we have investigated [7] a collapsar whose equation of state is an idealized form of neutron fluid*, and for which, above a certain mass, its maximum density lies between the event horizon and the photonsphere.

2 The exterior light orbits

Darwin [9, 10] described the null geodesics of the Schwarzschild metric

$$ds^2 = \frac{r-2m}{r}\,dt^2 - \frac{r}{r-2m}\,dr^2 - r^2 d\theta^2 - r^2\sin^2\theta\, d\phi^2, \quad (1)$$

where $2m$ is the gravitational radius. He extended the standard theory of light deflection, his method being equivalent to minimising the action integral for a light ray with small impact parameter starting from infinity; for orbits in the plane $\theta = \pi/2$,

$$\delta\int L d\phi = 0, \quad (2)$$

with the lagrangian

$$L = \left[\frac{r-2m}{r}\,t'^2 - \frac{r}{r-2m}\,r'^2 - r^2\right]^{1/2}, \quad (3)$$

where a prime denotes differentiation with respect to ϕ. The Lagrange equation for the cyclic coordinate t is

$$\left[\frac{d}{d\phi} - \frac{L'}{L}\right]\frac{r-2m}{r}\,t' = 0. \quad (4)$$

The corresponding conservation integral for ϕ enables us to put $L'/L = 2r'/r$, so we obtain

$$t' = \frac{r^3}{p(r-2m)}, \quad (5)$$

the constant p being the impact parameter

$$p = \lim_{r\to\infty} r^2\frac{d\phi}{dt}. \quad (6)$$

The ray orbit is then obtained by substituting for t' and then putting $L = 0$, that is

$$r'^2 = \frac{r^4}{p^2} - r^2 + 2mr. \quad (7)$$

Darwin deduced that a ray with impact parameter p greater than $3m\sqrt{3}$ returns to $r = \infty$; the deflection angle may be many multiples of 2π as p approaches $3m\sqrt{3}$, and in the limiting case $p = 3m\sqrt{3}$ the ray circles indefinitely at $r = 3m$, which is nowadays called the *photonsphere*. For p less than this, the ray is captured, and it goes to what Darwin termed the "barrier", nowadays called the *event horizon*, at $r = 2m$. He also repeated the point previously made by OS, that the journey from $r = 3m$ to $r = 2m$ takes an infinite time. When the collapse is incomplete, the surface being at $r = r_1 > 2m$, a ray arrives there making an angle with the normal of

$$\xi = \tan^{-1}\left[\left(\frac{r_1^2}{p^2} - 1 + \frac{2m}{r_1}\right)^{-1/2}\right], \quad (8)$$

and in the limiting case $r_1 = 2m$ this becomes

$$\xi = \tan^{-1}\left(\frac{p}{2m}\right). \quad (9)$$

We may deduce directly the orbits of rays exiting from the barrier; those falling within a cone of semiangle $\tan^{-1}(3\sqrt{3}/2) = 68.9$ degrees go to our telescope at "infinity", forming an image of parallax $6m\sqrt{3}$. Any collapsar with $2m < r_1 < 3m$ has this same parallax, but at $3m$ the cone has opened up fully to 90 degrees. A collapsar bigger than $3m$ has a parallax bigger than $6m\sqrt{3}$, while for much larger collapsars, like white dwarfs of solar mass, light deflection is negligible, and the parallax is then simply twice the surface radius. In Figure 1 a number of rays have been plotted, leaving various points in the surface, when $r_1 = 2.2m$, and going towards our telescope; we note that the rays going to the edge of the image come from points on the "invisible face" of the collapsar.

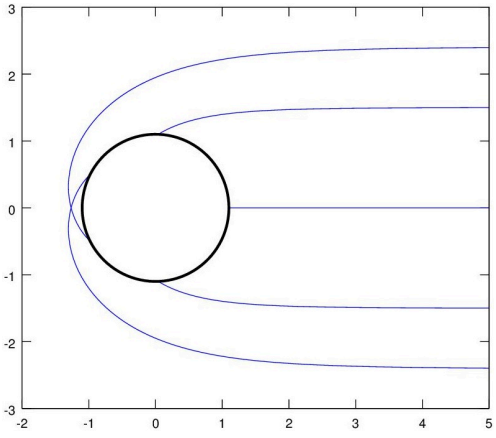

Fig. 1: The light rays issuing from the surface of a collapsar at 1.1 times the gravitational radius, collimated towards a distant telescope. The outer rays are close to the edge of the image, which has a diameter of 5.2 times the gravitational radius; these rays have their sources on what, in the absence of gravitational lensing, would be the invisible part of the surface. The unit of distance is the gravitational radius.

*This model is simply that of Oppenheimer and Volkoff [8] with a different boundary condition at the origin.

For Sagittarius A* the minimum parallax, according to the above analysis, and with the distance of EHT from the galactic centre equal to 2.4×10^{17} km, is 52 arc microseconds, which exceeds the best available current value [3] by about 50 percent. The image profile, that is its intensity $C(p)$ as p goes from zero to $3m\sqrt{3}$, is given by

$$C(p) = \left| \frac{r_1 \sin\phi_0 \cos\xi}{p} \frac{d\phi_0}{dp} \right| , \qquad (10)$$

where ξ is given by (8), that is

$$\cos\xi = \sqrt{\frac{r_1^3 - p^2 r_1 + 2p^2 m}{r_1^3 + 2p^2 m}} , \qquad (11)$$

and ϕ_0 is the angle between the outward normal at the surface and the ray's final direction, that is

$$\phi_0 = \int_{r_1}^{\infty} \frac{p dr}{\sqrt{r^4 - p^2 r^2 + 2mp^2 r}} , \qquad (12)$$

leading to

$$\frac{d\phi_0}{dp} = \int_{r_1}^{\infty} \frac{r^4 dr}{(r^4 - p^2 r^2 + 2mp^2 r)^{3/2}} . \qquad (13)$$

Note that, for $r_1 \gg 2m$, $\xi = \phi_0$, $p = r_1 \sin\phi_0$, and $C(p) = 1/r_1 = \text{const.}$ with ϕ_0 going from $-\pi/2$ to $\pi/2$, giving a uniform circular image of radius r_1; for our case ϕ_0 takes all real values. In Figure 2 the image profile $C(p)$ is plotted. The edge of the image is at $p = 3m\sqrt{3} = 2.598\, r_0$, where r_0 is the gravitational radius. Note that, though $C(p)$ drops to zero at $p = 2.388 r_0$, there is a bright fringe between that value and $p = 2.588\, r_0$; though not shown in the Figure, there is a series of narrower fringes between the latter value and the edge of the image at $p = 2.598\, r_0$. The fringes result from light rays circling close to the photonsphere before finally escaping to reach the telescope, their minima occurring at p-values for which ϕ_0 are integer multiples of π.

A ray which leaves the surface in a direction falling outside the limiting cone, that is with an orbit described by $p > 3m\sqrt{3}$, turns round before reaching the photonsphere, and returns to the barrier after an infinite time.

None of this accords with the OS description, in which the cone closes down to zero at $r = 2m$.

3 The interior light orbits

According to the OS [4] model, the surface of the collapsar completely contracts to the barrier only at $t = \infty$; in the words of that article

> ... an external observer sees the star asymptotically shrinking to its gravitational radius.

Specifically $r_1(t)$ is given by

$$t = -\frac{2}{3}\sqrt{\frac{r_1^3}{2m}} - 2\sqrt{2mr_1} + 2m \ln \frac{\sqrt{r_1} + \sqrt{2m}}{\sqrt{r_1} - \sqrt{2m}} . \qquad (14)$$

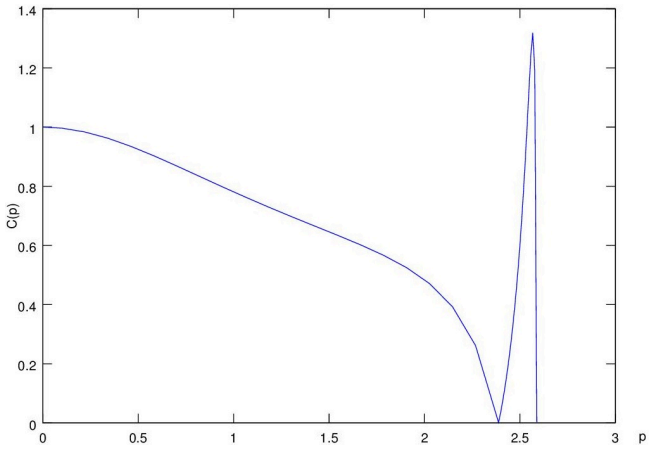

Fig. 2: The image profile $C(p)$ formed by the rays in Figure 1. Again the gravitational radius is the distance unit on the horizontal axis p, and $C(p)$ is normalized to $C(0) = 1$.

For $r < r_1$ the OS metric is

$$ds^2 = \frac{r^3}{2mR^3}\left(\frac{dr}{r} - \frac{dR}{R}\right)^2 - \frac{r^2}{R^2} dR^2 - \\ - r^2 d\theta^2 - r^2 \sin^2\theta\, d\phi^2 , \qquad (15)$$

where the coordinate R lies between 0 and 1, and is related to t in a manner to be determined by matching conditions imposed at the surface.

The interior null geodesics in the plane $\theta = \pi/2$ are constructed from the Lagrangian

$$L = \left[\frac{r^3}{2mR^3}\left(\frac{r'}{r} - \frac{R'}{R}\right)^2 - \frac{r^2}{R^2} R'^2 - r^2\right]^{1/2} . \qquad (16)$$

The Lagrange equations for r and R are, putting $L'/L = 2r'/r$ as in the exterior case,

$$\frac{2rr''}{R^3} - \frac{2r^2R''}{R^4} - \frac{3r'^2}{R^3} - \frac{2rr'R'}{R^4} + \frac{5r^2R'^2}{R^5} + \\ + \frac{4mrR'^2}{R^2} + 4mr = 0 \qquad (17)$$

and

$$\frac{2r^2r''}{R^4} - \frac{2r^3R''}{R^5} - \frac{3rr'^2}{R^4} - \frac{2r^2r'R'}{R^5} + \frac{5r^3R'^2}{R^6} + \\ + \frac{4mr^2R''}{R^2} - \frac{4mr^2R'^2}{R^3} = 0 . \qquad (18)$$

Combining these to eliminate r'', we obtain

$$R'' = R + \frac{2R'^2}{R} , \qquad (19)$$

for which a sufficiently general solution, for $0 < R < 1$, is

$$R = \sin\phi_0 \csc\phi \quad (\phi_0 < \phi < \pi - \phi_0) . \qquad (20)$$

Then, as in the exterior case, we obtain a first order equation for r by substituting this in L and putting $L = 0$, namely

$$r' = \sqrt{\frac{2mr\sin^3\phi_0}{\sin^5\phi}} - r\cot\phi\,, \tag{21}$$

with the solution

$$r = \frac{m\sin\phi_0}{2\sin\phi}\,(A - \sin\phi_0\cot\phi)^2. \tag{22}$$

A ray which arrives at $R = 1$, that is $\phi = \phi_0$, with $r = r_1$ has $A = 2\sqrt{r_1/(2m)} + \cos\phi_0$; the special case $\phi_0 = 0$ was given in eq (14) of [1]. At this point the ray has gradient

$$r' = \csc\phi_0\,\sqrt{2mr_1} - r_1\cot\phi_0\,. \tag{23}$$

4 Matching at the surface

OS [4] matched their metric with the exterior (1) by defining the *cotime* $y(r, R)$ related to t by

$$\frac{t}{2m} = -\frac{2}{3}\,y^{3/2} - 2\sqrt{y} + \ln\frac{\sqrt{y}+1}{\sqrt{y}-1}\,; \tag{24}$$

this they required to satisfy $y(r_1, 1) = r_1/(2m)$. To match the two metric tensors at $R = 1$ they then put

$$y = \frac{r}{2Rm} + \frac{R^2-1}{2}\,. \tag{25}$$

With the metrics matched at the surface, that means the refractive indices are also matched, so the corresponding light rays should join smoothly there. Eq (23) for the interior ray gives, at $r = r_1$,

$$r'^2 + r^2 - 2mr = \left(r_1\csc\phi_0 - \sqrt{2mr_1}\cot\phi_0\right)^2, \tag{26}$$

so the value

$$p = \frac{r_1^2\sin\phi_0}{r_1 - \sqrt{2mr_1}\cos\phi_0} \tag{27}$$

gives a smooth connection between the interior (23) and exterior (7) rays at $r = r_1$. Differentiating (24) and (25), we then find that the values of t' also match at r_1, which confirms that the light speed r'/t' is continuous there.

It may now be seen that, as r_1 approaches $2m$, the speed of light at the surface goes to zero, which generalizes the particular case treated in [1], where the light ray was normal to the surface. Such behaviour may be understood as resulting from the infinite "dust" density there (see below). This behaviour will be modified by the intervention of nongravitational forces; in particular we have studied the effect of the Fermi degeneracy pressure in a neutron star [7], for which the density has a finite maximum well separated from both the surface and $r = 0$. Thus, for a collapsar made of real stellar matter, it makes sense to consider a state of equilibrium whose radius exceeds the gravitational, and for which light leaves the surface with a finite speed; this was the situation depicted in Figure 1.

I add that the matching relation (25) is not unique, though OS stated that it was. In my previous articles [1, 2] the alternative

$$y = \frac{r}{2Rm} - \frac{(1-R)(5-R)}{4} \tag{28}$$

was given. This is part of a wider family of matching relations, and, for this particular choice, has certain advantages in respect of causality.

The infinite surface density of the OS final state may be seen in their calculation of the scalar density ρ, namely

$$\rho = \frac{3R^3}{8\pi r^3}\,. \tag{29}$$

Multiplying this by their three-volume element, we obtain

$$\rho\sqrt{-g}\,dR\,d\theta\,d\phi = \frac{3R^2\sin\theta}{8\pi}\,dR\,d\theta\,d\phi\,, \tag{30}$$

which, in terms of r, gives the density

$$\rho\sqrt{-g}\,dr\,d\theta\,d\phi = \frac{3R^2\sin\theta}{8\pi}\left(\frac{\partial R}{\partial r}\right)_t dr\,d\theta\,d\phi\,. \tag{31}$$

The partial derivative is given, at cotime $y = 1$, by

$$\left(\frac{\partial R}{\partial r}\right)_t = \left(\frac{\partial R}{\partial r}\right)_y = \frac{1}{3m(1-R^2)}\,, \tag{32}$$

giving infinite density at $R = 1$.

Actually we have found that the density in the shell just inside r_1 is very much reduced for a supermassive object like Sagittarius A*, and I propose that the material there is an electron gas with a nearly stationary nucleonic background* which should have broadly similar optical properties to both the OS dust cloud and the neutron star. In these cases the light speed will still be considerably reduced near the surface, but will remain finite.

5 Discussion

The suggestion about the origin of the EHT image of Sagittarius A*, namely that part of the light we receive comes from the collapsar itself, has implications for the direction future observations with the telescope should take. A central problem is to explain the present-day value of the parallax, which is 37μas as opposed to the 52μas we obtained in Section 2. We note that the size of this image is not at all well defined, because of the need to separate the signal from the background noise of nearby objects; this is reflected by the wide error bar in the above parallax. It should be noted also that the image of the accretion disc has almost the same diameter as the one

This entails classifying Sagittarius A as a *supermassive white giant*.

described in Section 2 if the disc is inside the photonsphere, and that its image is larger if it lies outside the photonsphere.

The fringes of the image, described in Section 2, do not seem to have been noticed previously, though they are surely present also in the image of the accretion disc. To distinguish between the two images, arising, as they do, from two superimposed sources of almost the same diameter, will require further analysis along the lines of Section 2; the principal difference is the three-dimensional form of the collapsar, as opposed to the flat, effectively two-dimensional form of the disc. Some progress, both in image enhancement and in theoretical modelling, would help to clarify matters.

The classic article of Oppenheimer and Snyder [4], based in turn on the equally classic one of Tolman [11], was essential for the construction of the matched orbits. In particular these articles (see also [12]) enable us to identify the comoving coordinate R used in Section 3. But the step required to describe fully the orbits of particles of "dust", that is the stellar material, and of light rays near the surface, is the identification of the time coordinate $t(r, R)$ made in our earlier article [1] and in Section 4 of this one.

Acknowledgement

I gratefully acknowledge the help of my colleagues Clifford Chafin, Charles Kinniburgh, Robin Spivey and Max Wallis.

Submitted on May 1, 2016 / Accepted on May 4, 2016

References

1. Marshall T.W. *Astrophys. Space Sci.*, 2012, v.342, 329–332.
2. Marshall T.W. *Progress in Physics*, 2016, v.12, issue 3, 219–221.
3. Doeleman S.S., Weintroub J.W., Rogers A.E.E., Plambeck R., Freund R., Tilanus R., Ziurys L., Moran J., Corey B., Young K.H., Smythe D.L., Titus M., Marrone D.P., Cappallo R.J., Bock D.C.J., Bower G.C., Chamberlin R., Maness H., Niell A.E., Roy A., Strittmatter P., Werhimmer D., Whitney A.R. and Woody D. *Nature*, 2008, v.455, 78.
4. Oppenheimer J.R. and Snyder H. *Phys. Rev.*, 1939, v.56, 455.
5. Penrose R. *Phys. Rev. Letters*, 1965, v.14, 57.
6. Thorne K.S. Black Holes and Time Warps. Norton, New York, 1994, see *Prologue*.
7. Wallis M.K. and Marshall T.W. Energy in General Relativity – the case of the neutron star. *Physical Interpretations of Relativity Theory*, Proceedings of International Meeting, Bauman State Technical University, Moscow, 2015 pp. 544–556 (2015). See also http://tinyurl.com/PIRT-Moscow-15-MarshallWallis
8. Oppenheimer J.R. and Volkoff G.M. *Phys. Rev.*, 1939, v.54, 540.
9. Darwin, Sir Charles. *Proc. Roy. Soc. A*, 1958, v.249, 180–194.
10. Darwin, Sir Charles. *Proc. Roy. Soc. A*, 1961, v.263, 39–50.
11. Tolman R.C., *Proc. Nat. Acad. Sci. (USA)*, 1934, v.20, 169–176.
12. Weinberg S. Gravitation and Cosmology. John Wiley, New York, 1972, pp. 338–341.

The Dual Behavior of Quantum Fields and the Big Bang

Malik Matwi

Damascus University, Damascus, Syria. E-mail: malik.matwi@hotmail.com

We modify the propagator of the quantum fields for the quarks and gluons. With that we have finite results (without ultraviolet divergence) in the perturbation theory. Then we search for $a^2p^2 \to 0$ and $a^2k^2 \to 0$ with fixing the Lagrangian parameters Z_i, therefore we can ignore our modification. We find the situation $a^2p^2 \to 0$ and $a^2k^2 \to 0$ associates with the free particles situation $g \to 0$ (g is the coupling constant) and the situation $a \neq 0$ associates with the perturbation breaking. We try to give the modification terms $a^2p^2/(1+a^2p^2)$ and $a^2k^2/(1+a^2k^2)$ physical aspects, for that we find the corresponding terms in the Lagrangian. To do that we find the role of those terms in the Feynman diagrams, in self energies, quarks gluons vertex, ... We see we can relate the propagator modification to fields dual behavior, pairing particle-antiparticle appears as scalar particles with mass $1/a$. For the quarks we can interrupt these particles as pions with charges $(-1, 0, +1)$. If we used the propagator modification for deriving the quarks static potential $U(r)$ of exchanged gluons and pions we find $U(0) \sim 1/a$ if we compare this with the Coulomb potential we find the length a equivalent to the smallest distance between the interacting quarks. We use the static potential in quarks plasma study. We find the free and confinement quarks phases. We suggest a nuclear compression. We find there is a decrease in the global pressure due to the nuclear condensation. We use this decrease in the Friedman equations solutions, we find we can control the dark matter and dark energy, we can cancel them.

1 Quarks and gluons propagator modification

To remove the ultraviolet (UV) divergences in the quarks and gluons perturbed interaction, we modify the propagator like:

$$\overline{\Delta}^{ab}_{\mu\nu}(k^2) = \frac{g_{\mu\nu}\,\delta^{ab}}{k^2 - i\varepsilon}\left(1 - \frac{a^2k^2}{1+a^2k^2}\right) \quad \text{for gluons} \tag{1.1}$$

$$\bar{S}_{ij}(\not{p}) = \frac{-\not{p}\delta_{ij}}{p^2 - i\varepsilon}\left(1 - \frac{a^2p^2}{1+a^2p^2}\right) \quad \text{for quarks} \tag{1.2}$$

the indexes a and b are gluons indexes, i and j color indexes and a is critical length, $\hbar = c = 1$. We use this modification in calculating the quarks self-energy for the perturbation interaction with the gluons, then we renormalize the interaction and search for the condition $a^2p^2 \to 0$ and $a^2k^2 \to 0$. We have

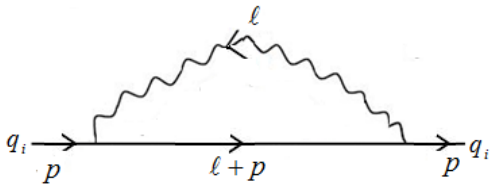

Fig. 1: The quarks self energy in strong interaction.

$$i\Sigma_{ij}(\not{p}) = \int \frac{d^4\ell}{(2\pi)^4}\left[ig_s\gamma^\mu T^a_{ik}\frac{\overline{S}_{kl}(\not{p}+\not{\ell})}{i}ig_s\gamma^\nu T^b_{lj}\right]\frac{\overline{\Delta}^{ab}_{\mu\nu}(\ell^2)}{i}$$
$$= g_s^2 T^a_{ik}T^b_{lj}\int \frac{d^4\ell}{(2\pi)^4}\left[\gamma^\mu\frac{(-\not{p}-\not{\ell})\delta_{kl}}{(p+\ell)^2}\gamma^\nu\right]\frac{g_{\mu\nu}\delta^{ab}}{\ell^2}\,.$$

So

$$i\Sigma_{ij}(\not{p}) = g_s^2 T^a_{ik}T^a_{kj}\int \frac{d^4\ell}{(2\pi)^4}\left[\gamma^\mu\frac{(-\not{p}-\not{\ell})}{(p+\ell)^2}\gamma^\nu\right]\frac{g_{\mu\nu}}{\ell^2}$$
$$= g_s^2 C(R)\delta_{ij}\int \frac{d^4\ell}{(2\pi)^4}\left[\gamma^\mu\frac{(-\not{p}-\not{\ell})}{(p+\ell)^2}\gamma_\mu\right]\frac{1}{\ell^2}$$

using $\gamma^\mu(-\not{p}-\not{\ell})\gamma_\mu = 2(-\not{p}-\not{\ell})$, it becomes

$$i\Sigma_{ij}(\not{p}) = 2g_s^2 C(R)\delta_{ij}\int \frac{d^4\ell}{(2\pi)^4}\frac{(-\not{p}-\not{\ell})}{(p+\ell)^2}\frac{1}{\ell^2}\,.$$

Now we use the gluon modified propagator

$$\overline{\Delta}^{ab}_{\mu\nu}(k^2) = \frac{g_{\mu\nu}\delta^{ab}}{k^2 - i\varepsilon}\left(1 - \frac{a^2k^2}{1+a^2k^2}\right)$$

we get

$$i\Sigma_{ij}(\not{p}) = 2g_s^2 C(R)\delta_{ij}$$
$$\int \frac{d^4\ell}{(2\pi)^4}\frac{(-\not{p}-\not{\ell})}{(p+\ell)^2}\frac{1}{\ell^2}\left(1 - \frac{a^2\ell^2}{1+a^2\ell^2}\right) \tag{1.3}$$
$$= 2g_s^2 C(R)\delta_{ij}\int \frac{d^4\ell}{(2\pi)^4}\frac{(-\not{p}-\not{\ell})}{(p+\ell)^2}\frac{1}{\ell^2}\frac{1}{1+a^2\ell^2}\,. \tag{1.4}$$

For massive quarks, the self-energy becomes:

$$i\Sigma_{ij}(\not{p}) = g_s^2 C(R)\delta_{ij}\int \frac{d^4\ell}{(2\pi)^4}\frac{N}{(p+\ell)^2+m_q^2}\frac{1}{\ell^2+m_\gamma^2}\frac{1}{1+a^2\ell^2}$$

with $N = \gamma^\mu(-\not{p} - \not{\ell} + m)\gamma_\mu$, using the Feynman formula:

$$\frac{1}{((p+\ell)^2+m^2)\cdot(\ell^2+m_\gamma^2)\cdot(1/a^2+\ell^2)} = \int dF_3 \frac{1}{\left[((p+\ell)^2+m^2)x_1 + (\ell^2+m_\gamma^2)x_2 + (1/a^2+\ell^2)x_3\right]^3}$$

with $\int dF_3 = 2\int_0^1 dx_1 dx_2 dx_3\, \delta(x_1+x_2+x_3-1)$ and setting the transformation $q = \ell + x_1 p$ with changing the integral to be over q and making transformation to Euclidean space, the self-energy becomes [2]

$$i\Sigma_{ij}(\not{p}) = g_s^2 C(R)\delta_{ij} i \int \frac{d^4\bar{q}}{(2\pi)^4}\frac{1}{a^2}\int dF_3 \frac{N}{[\bar{q}^2+D]^3}$$

with $D = -x_1^2 p^2 + x_1 p^2 + x_1 m^2 + x_2 + m_\gamma^2 + (1-x_1-x_2)1/a^2$. The linear term in q integrates to zero, using $q = \ell + x_1 p$, N is replaced with [2]

$$N \to -2(1-x_1)\not{p} - 4m\,.$$

Using the relation

$$\int \frac{d^d\bar{q}}{(2\pi)^d}\frac{(\bar{q}^2)^a}{(\bar{q}^2+D)^{\mathrm{b}}} = \frac{\Gamma(\mathrm{b}-a-\frac{d}{2})\Gamma(a+\frac{d}{2})}{(4\pi)^{\frac{d}{2}}\Gamma(\mathrm{b})\Gamma(\frac{d}{2})} D^{-(\mathrm{b}-a-\frac{d}{2})}\,,$$

the integral over q in Euclidean space becomes:

$$\Sigma_{ij}(\not{p}) = g_s^2 C(R)\delta_{ij}\frac{1}{a^2}\int dF_3 N \frac{\Gamma(3-2)\Gamma(2)}{(4\pi)^2\Gamma(3)\Gamma(2)} D^{-(3-2)} = g_s^2 C(R)\delta_{ij}\frac{1}{a^2}\int dF_3 \frac{N}{16\pi^2\times 2} D^{-1}\,.$$

The self-energy becomes

$$\Sigma_{ij}(\not{p}) = g_s^2 C(R)\delta_{ij}\frac{1}{a^2}\int_0^1 dx_1\int_0^{1-x_1} dx_2 \frac{N}{16\pi^2}\frac{1}{D} = \frac{g_s^2 C(R)\delta_{ij}}{16\pi^2}\int_0^1 dx_1\int_0^{1-x_1} dx_2 \frac{-2(1-x_1)\not{p} - 4m}{a^2\left[-x_1^2 p^2 + x_1 p^2 + x_1 m^2 + x_2 m_\gamma^2 + (1-x_1-x_2)/a^2\right]}\,.$$

We write

$$\Sigma_{ij}(\not{p}) = C(R)\delta_{ij}\frac{g_s^2}{8\pi^2}\int_0^1 dx_1\int_0^{1-x_1} dx_2 \frac{-(1-x_1)\not{p} - 2m}{[a^2 f + (1-x_1-x_2)]} \tag{1.5}$$

with $f = -x_1^2 p^2 + x_1 p^2 + x_1 m^2 + x_2 m_\gamma^2$ this is a finite result (without divergences).

Now we renormalize the fermions propagator to give the real states and let $a \to 0$. The interacting quarks propagator becomes [2]:

$$\overline{S}(\not{p})^{-1} = \not{p} + m - \Sigma(\not{p})\,. \tag{1.6}$$

To renormalize the interacting field, we write it as

$$\overline{S}(\not{p})^{-1} = \not{p} + m - \Sigma(\not{p}) = Z_2\not{p} + Z_m m\,. \tag{1.7}$$

The parameters Z_2 and Z_m are the renormalization parameters, later we try to make them constants. For the interacting field ψ we have:

$$\langle 0|\,\psi(\not{p})\bar{\psi}(-\not{p})\,|0\rangle = \frac{1}{i}\frac{1}{\not{p}+m-\Sigma(\not{p})} = \frac{1}{i}\frac{1}{Z_2\not{p}+Z_m m} = \frac{1}{iZ_2}\frac{1}{\not{p}+Z_2^{-1}Z_m m}\,.$$

We can rewrite as

$$\langle 0|\,\sqrt{Z_2}\psi(\not{p})\sqrt{Z_2}\bar{\psi}(-\not{p})\,|0\rangle = \frac{1}{i}\frac{1}{\not{p}+Z_2^{-1}Z_m m}$$

and make $m_0 = Z_2^{-1}Z_m m$ and $\psi_0 = \sqrt{Z_2}\psi$ with that we have bare fields ψ_0 that are like the free fields and like the classical fields, so we can make them independent of the interaction, so $\partial\psi_0/\partial p^2 = \partial m_0/\partial p^2 = 0$ for $a \to 0$ and by that we renormalize the interaction. We make ψ the interacting field with mass m the physical mass, but we have to make $\Re\,[\Sigma(-m)] = 0$ in (1.6) but with $m_\gamma^2 < 0$. From (1.5) and (1.7) we have

$$Z_2 = 1 + C(R)\frac{g_s^2}{8\pi^2}\int_0^1 dx_1\int_0^{1-x_1} dx_2 \frac{1-x_1}{[a^2 f + (1-x_1-x_2)]}$$

$$Z_m = 1 + C(R)\frac{g_s^2}{8\pi^2}\int_0^1 dx_1\int_0^{1-x_1} dx_2 \frac{2}{[a^2 f + (1-x_1-x_2)]}$$

and $f = -x_1^2 p^2 + x_1 p^2 + x_1 m_q^2 + x_2 m_\gamma^2$.

By that we remove the self-energy of the interacting quark and make the mass variable. For easiness we ignore m_q and m_γ so

$$Z_2 = 1 + C(R)\frac{\alpha_s}{2\pi}\int_0^1 (1-x)\ln\left(1+\frac{1}{a^2p^2x}\right)dx = 1 + \frac{C(R)\alpha_s}{4\pi(a^2p^2)^2}\left[(a^2p^2)^2\ln\left(1+\frac{1}{a^2p^2}\right) - a^2p^2 + (2a^2p^2+1)\ln(a^2p^2+1)\right].$$

Now we fix $Z_2 = constant$ and search for the situations $-a^2p^2 \to 0$ for timelike and $a^2p^2 \to 0$ for spacelike, we have

$$\frac{\alpha_s}{(a^2p^2)^2}\left[(a^2p^2)^2\ln\left(1+\frac{1}{a^2p^2}\right) - a^2p^2 + (2a^2p^2+1)\ln\left(a^2p^2+1\right)\right] = \mathrm{c}\,.$$

For spacelike $p^2 > 0$, we have Fig. 2. According to this figure, we have $a^2p^2 = \exp(-c/\alpha_s) \to 0$ when $\alpha_s \to 0$ this is the

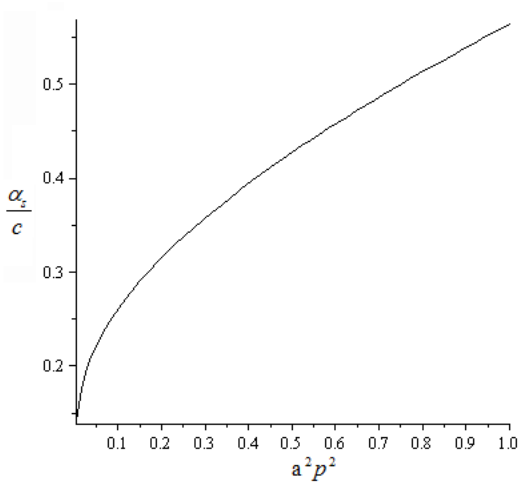

Fig. 2: The behavior of the length a with fixing Z_2.

decoupling; $p^2 \gg \Lambda^2_{QCD}$. It is the free quarks and gluons situation; $\alpha_s \to 0$ occurs at high energy for the free quarks phase. Because $ap \to 0$ so $p \ll 1/a$ this gives $r \gg a \to 0$, therefore the propagator modification is ignored. So the behavior of the length a is like the behavior of the coupling constant α_s and the modification terms are removed $ap \ll 1$ at high energy (free quarks phase).

For the limited low energy we fix $\alpha_s/a^2 = constant \times \sigma$, σ is string tension that appears in the low energy static potential $U(r)$ as we will see, for $a \to 0$ we have

$$Z_2 = 1+C(R)\frac{\alpha_s}{4\pi}\left(\frac{3}{2} - \ln\left(p^2a^2\right) + O(p^2a^2)\right) \to 1, \text{when } a \to 0$$

$$Z_m = 1 + C(R)\frac{\alpha_s}{\pi}\left(1 - \ln\left(p^2a^2\right) + O(p^2a^2)\right) \to 1, \text{when } a \to 0$$

We know the strong interaction coupling constant α_s increases extremely at low limited energy, therefore, according to the figure, we can't let $a \to 0$, so we assume when the perturbation breaks down the length a could not be removed and takes non-zero value, let it be a_0, so the propagator modification takes place.

1.1 The confinement situation

According to Fig. 2 it is possible to have $ap > 1$ (the coupling constant α_s increases extremely at low energy), therefore $p > 1/a \to r < a$ which is the quarks confinement phase at low energy.

To study the quarks confinement, we use the modified gluons propagator in deriving the static potential of the quark-quark gluons exchange. We define this potential in momentum space using M matrix elements for quark-quark (gluons exchange) interaction, with $\omega_0 = k_0 = 0$ (like the Born approximation to the scattering amplitude in non-relativistic quantum mechanics [1])

$$iM = -i\tilde{V}(k)J^\mu(p'_2, p_2)J_\mu(p'_1, p_1)$$

with the transferred current $J^\mu(p', p) = \bar{u}(p')\gamma^\mu u(p)$ where spinor states $u(p)$ include the helicity states.

We find M matrix elements using the Feynman diagrams for quark-quark gluons exchange using color representation for one quark like

$$u(p)_{color\otimes spinor} = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}u(p)_{spinor}\,.$$

For distinguishable quarks (only one diagram), we have

$$iM = \bar{u}^i(p'_2)ig_s\gamma^\mu(T^a)^j_i u_j(p_2)\frac{\Delta^{ab}_{\mu\nu}(k^2)}{i}\bar{u}^k(p'_1)ig_s\gamma^\nu(T^b)^\ell_k u_\ell(p_1)$$

with $k = p'_2 - p_2 = p_1 - p'_1$.

Using Gell-Mann matrices, we consider the matrices $T^a = \lambda^a; \lambda_1, \ldots, \lambda_8$ as $SU(3)$ generators, and using the modified gluons propagator we have

$$iM = \sum_{ijk\ell} ig_s^2\bar{u}^i(p'_2)\gamma^\mu(T^a)^j_i u_j(p_2)\,\frac{g_{\mu\nu}\delta^{ab}}{k^2}\left(1 - \frac{a^2k^2}{1+a^2k^2}\right)$$
$$\bar{u}^k(p'_1)\gamma^\nu(T^b)^\ell_k u_\ell(p_1)$$

to sum over the color indexes i, j with the color representation like above and over gluon index a we write

$$\sum_{ij}\bar{u}^i(p'_2)\gamma^\mu(T^a)^j_i u_j(p_2)$$
$$= \bar{u}(p'_2)\gamma^\mu\,\frac{1}{\sqrt{3}}\begin{pmatrix}1 & 1 & 1\end{pmatrix}(T^a)\,\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}u(p_2)$$

and

$$\frac{1}{\sqrt{3}}\begin{pmatrix}1 & 1 & 1\end{pmatrix}(T^a)\,\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix} = \frac{1}{3}\sum_{ij}(T^a)^j_i.$$

Therefore the M matrix elements become

$$M = \frac{1}{9}\sum_a\left(\sum_{ij}(T^a)^j_i\right)^2 g_s^2\bar{u}(p'_2)\gamma^\mu u(p_2)\,\frac{1}{k^2}\left(1 - \frac{a^2k^2}{1+a^2k^2}\right)$$
$$\bar{u}(p'_1)\gamma_\mu u(p_1)\,.$$

The Gell-Mann matrices with nonzero sum of the elements are

$$\lambda_1 = \begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix},\ \lambda_4 = \begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix} \text{ and } \lambda_6 = \begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}.$$

So

$$\sum_a \left(\sum_{ij} (T^a)_i^j \right)^2 = 3\,(2)^2 = 12.$$

Therefore we have

$$M = \frac{12 g_s^2}{9} \frac{1}{k^2} \left(1 - \frac{a^2 k^2}{1 + a^2 k^2} \right) \bar{u}(p_2') \gamma^\mu u(p_2) \bar{u}(p_1') \gamma_\mu u(p_1)\,.$$

We have the potential $\tilde{V}(k)$ in momentum space as we defined

$$\begin{aligned} iM &= -i\tilde{V}(k) J^\mu(p_2', p_2) J_\mu(p_1', p_1) \\ &= i \frac{12 g_s^2}{9} g_s^2 \bar{u}(p_2') \gamma^\mu u(p_2) \frac{1}{k^2} \left(1 - \frac{k^2}{k^2 + 1/a^2} \right) \bar{u}(p_1') \gamma_\mu u(p_1) \end{aligned}$$

with the transferred currents $J^\mu(p_2', p_2) = \bar{u}(p_2') \gamma^\mu u(p_2)$ and $J^\mu(p_1', p_1) = \bar{u}(p_1') \gamma^\mu u(p_1)$. So we have

$$\tilde{V}(k) = -\frac{4 g_s^2}{3} \frac{1}{k^2} \left(1 - \frac{k^2}{k^2 + 1/a^2} \right).$$

Making the Fourier transformation to the space XYZ, we have the static potential $U(x)$ ($k_0 = 0$) like the electric potential [1]

$$\begin{aligned} U(x) &= \int \frac{d^3 k}{(2\pi)^3} \tilde{V}(k)\, e^{ik \cdot x} \\ &= -\frac{4 g_s^2}{3} \int \frac{d^3 k}{(2\pi)^3} \frac{1}{k^2} \left(1 - \frac{k^2}{k^2 + 1/a^2} \right) e^{ik \cdot x} \\ &= -\frac{4 g_s^2}{3 \times 4\pi r} \left(1 - \exp\left(-\frac{r}{a} \right) \right) \quad \text{with } r = \sqrt{x^2 + y^2 + z^2}\,. \end{aligned}$$

For low limited energy we have $ap > 1$ (Fig. 2) so $r < a$, the static potential becomes

$$U(r) = -\frac{4 g_s^2}{3 \times 4\pi r} \left[1 - \exp\left(-\frac{r}{a} \right) \right] = -u_0 + a_1 r - a_2 r^2 + \ldots$$

with

$$\begin{aligned} u_0 &= \frac{4}{3} \frac{g_s^2}{4\pi a} = \frac{4\alpha_s}{3a}, \\ a_1 &= \sigma = \frac{g_s^2}{3 \times 2\pi a^2} = \frac{2\alpha_s}{3a^2}, \\ a_2 &= \frac{4\alpha_s}{3 \times 6a^3}. \end{aligned}$$

To fix $u_0 = 4\alpha_s/3a$ we write it as

$$u_0 = \frac{4\alpha_s}{3a} = \frac{4\alpha_s}{3a^2} a = 2\sigma a$$

fixing the string tension σ and the length $a \to a_0$ at low energy.

This potential appears at low limited energy and prevents the quarks from spreading away, $r < a$ so it holds the quarks inside the hadrons. But starting from the high energies $a \to 0$, although the quarks masses are small but they are created only at high energies where they are free and by dropping the energy the situation $r < a$ appears, the length a would run and becomes higher at low energies, so have $-a^2 k^2 > 1$ for $r < a$ which is the confinement. The confinement (at low limited energy) means when $r \to a$ the two interacting quarks kinetic energy becomes zero (ignore the quark mass), therefore the highest kinetic energy that the quark can get equals σa which relates to the potential $U(r) = -u_0 + \sigma r + \ldots$ for $r < a$.

We can make $U(r)$ the potential for all quarks in $r < a$ so $\sigma \to \sum \sigma$ and consider r as average distance between the interacting quarks, so the energy σa becomes the highest kinetic energy of all quarks. When $r \to a$ the potential becomes $U(0) = -u_0 = -4\alpha_s/3a = -\sigma a < 0$ therefore the total quarks energy becomes negative.

In this situation the free quarks disappear, they become condensed in the hadrons. So the role of the potential is reducing the number of free quarks. Therefore the potential $u_0 = \sigma a$ leads to decrease of the free quarks chemical potential μ_0, and we have

$$\begin{aligned} \mu_0 \to \mu_0 + U(r) &= \mu_0 - \frac{\alpha_s}{r} \left(1 - e^{-r/a} \right) = \mu(r) \\ &\approx \mu_0 - u_0 + \sigma r \quad \text{for } r < a \end{aligned}$$

where we replaced $4\alpha_s/3$ with α_s. We renormalize this step at high energy for the free quarks, quarks plasma.

2 The quarks field dual behavior

To have finite results in the perturbation interaction, we modified the propagator like

$$\overline{\Delta}_{\mu\nu}^{ab}(k^2) = \frac{g_{\mu\nu} \delta^{ab}}{k^2 - i\varepsilon} \left(1 - \frac{a^2 k^2}{1 + a^2 k^2} \right) \text{ for gluons}$$

$$\bar{S}_{ij}(\not{p}) = \frac{-\not{p} \delta_{ij}}{p^2 - i\varepsilon} \left(1 - \frac{a^2 p^2}{1 + a^2 p^2} \right) \text{ for quarks}\,.$$

We saw we can ignore the modification terms $a^2 p^2/(1 + a^2 p^2)$ and $a^2 k^2/(1 + a^2 k^2)$ at high energy, but when the energy drops down to limited energy, those terms take place, we can give them a physical meaning, for that we search for the corresponding terms in the Lagrangian.

To do this, we find the role of those terms in the Feynman diagrams, in self energies, quarks-gluons vertex, ... We find that the terms $a^2 p^2/(1 + a^2 p^2)$ and $a^2 k^2/(1 + a^2 k^2)$ can be related to pairing quark-antiquark that appear as scalar particles with mass $1/a$ and charges $(-1, 0, +1)$ and we can interpret these particles as pions.

That appears in the particles-antiparticles composition in Feynman diagrams which mean for the fields, there is fields dual behavior, free fields and composite fields, this behavior leads to the possibility of separating the particles and possibility for their composition, so the dual behavior of the fields is elementary behavior. In general, for any particle A and its

antiparticle $\overline{A}$, in pertubation interaction, they pair and have a scalar particle $\overline{A}A$, this leads to reduce the currents (charges) of particles and antiparticles.

That is, for each outcoming particle, in Feynman diagrams, there is incoming antiparticle with positive energy and negative mass, depending on the coupling constant behavior (this is at high energy for the electromagnetic interaction and at low energy for the strong interaction, quarks and gluons). Therefore reducing their interactions with the charges in a way leads to finite results in the perturbation results.

Using the gluons modified propagator, the quark self-energy becomes (1.3)

$$i\Sigma_{ij}(\not p) = 2g_s^2 C(R)\delta_{ij}\int \frac{d^4\ell}{(2\pi)^4}\,\frac{(-\not p-\not\ell)}{(p+\ell)^2}\,\frac{1}{\ell^2}\left(1-\frac{a^2\ell^2}{1+a^2\ell^2}\right).$$

We can separate it into two parts

1. Quark−gluon part:

$$i\Sigma_{ij}(\not p) = 2g_s^2 C(R)\delta_{ij}\int \frac{d^4\ell}{(2\pi)^4}\,\frac{(-\not p-\not\ell)}{(p+\ell)^2}\,\frac{1}{\ell^2}\,;$$

2. pairing quarks part:

$$i\Sigma_{ij}(\not p) = 2g_s^2 C(R)\int \frac{d^4\ell}{(2\pi)^4}\,\frac{(-\not p-\not\ell)\delta_{ij}}{(p+\ell)^2}\,\frac{1}{\ell^2}\left(-\frac{a^2\ell^2}{1+a^2\ell^2}\right).$$

It appears that in the pairing part there is a scalar field φ propagator:

$$\frac{1}{i}\,\frac{1}{\ell^2+1/a^2}$$

which is real scalar particles field propagator with mass $1/a$, to preserve the charges, spin, ..., this particle must be condensed of quark-antiquark $|\overline{q}q\rangle$ (particle-antiparticle in general) so we have new diagram (Fig. 3), we rewrite

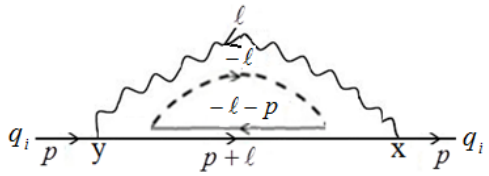

Fig. 3: Representation the dual behavior, joined particle−antiparticle with opposite momentum−energy.

$$i\Sigma_{ij}(\not p) = 2(-g_s)^2 C(R)\int \frac{d^4\ell}{(2\pi)^4}\,\frac{(-\not p-\not\ell)\delta_{ij}}{i(p+\ell)^2}\,\frac{-i}{(\ell)^2+1/a^2}\,.$$

Therefore we must add new interaction terms to the quarks Lagrangian, the possible terms are:

$$\Delta L = -ig_{\varphi q}\varphi\bar{Q}Q \quad \text{with} \quad g_{\varphi q} = g_s\sqrt{2C(R)}$$

or

$$\Delta L = g_{\varphi q}\varphi\bar{Q}\gamma_5 Q\,.$$

We expect the pairing particles-antiparticles preserve the flavor symmetry, so the real scalar field φ becomes $|\overline{q_i}q_j\rangle$. For two flavors q_i and q_j we write the quarks field like $Q = \begin{pmatrix} q_i & q_j \end{pmatrix}^T$ so

$$\Delta L = -ig_{\varphi q}\varphi^a\bar{Q}T_2^a Q \quad \text{or} \quad \Delta L = g_{\varphi q}\varphi^a\bar{Q}T_2^a\gamma_5 Q\,.$$

The real scalar fields φ^a could interact with itself and have real non-zero ground value υ then $\langle\varphi\rangle = \upsilon$ so we can renormalize it like

$$\varphi^a T_2^a \to \nu - i\nu\pi^a T_2^a + ...$$

then we have

$$\begin{aligned}\Delta L &= -ig_{\varphi q}\bar{Q}\left(\nu - i\nu\pi^a T_2^a + ...\right)Q\\ &= -ig_{\varphi q}\nu\bar{Q}Q - g_{\varphi\pi}\pi\bar{Q}Q + ...\ \text{Chiral symmetry breaking}\end{aligned}$$

or

$$\Delta L = g_{\varphi q}\bar{Q}(\nu - i\nu\pi^a T_2^a + ...)\gamma_5 Q \to g_{\pi q}\bar{Q}\gamma_5 Q - ig_{\pi q}\pi\bar{Q}\gamma_5 Q + ...$$

Here the particles $\pi^a T_2^a \to \pi = (\pi^0, \pi^-, \pi^+)$ are the pions. The unusual terms $-ig_{\varphi a}\nu\bar{Q}Q$ and $g_{\pi q}\bar{Q}\gamma_5 Q$ are not hermitian and violate the symmetries, so they let the quarks disappear, damping at low energy $r < a$:

$$\begin{aligned}e^{i\Delta Et}\,|Q\rangle &= e^{-i\Delta Lt}\,|Q\rangle = e^{-g_{\varphi q}\nu\bar{q}qt}\,|Q\rangle\\ &= \sum_n e^{-g_{\varphi q}\nu(\bar{q}q)t}\,|E_n\rangle\langle E_n\,|Q\rangle \to |0\rangle\langle 0\,|Q\rangle\,.\end{aligned}$$

E_n is the energy of the quarks in state $|n\rangle$ and $e^{i\hat{H}t}\,|Q\rangle$ is the eigenstate of the quarks field operator $\hat{Q}(t)$ in Heisenberg picture, $\hat{Q}(t) = e^{i\hat{H}t}\hat{Q}e^{-i\hat{H}t}$.

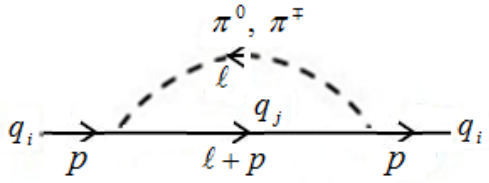

Fig. 4: The quarks interaction with pions as a result of dual behavior.

That damping in the states is because of the pairing quark-antiquark at low energy $a \neq 0$, this pairing reduces the charges (currents) of free quarks (Fig. 5). We can see that if we relate the minus sign in $-a^2\ell^2/(1+a^2\ell^2)$ to the fermions propagator:

$$S(x-y) = \int \frac{d^4p}{(2\pi)^4}\,\frac{-\not p}{p^2}e^{ip(x-y)}\ \text{(propagator from } y \text{ to } x)$$

so

$$-S(x-y) = -\int \frac{d^4p}{(2\pi)^4}\,\frac{-\not p}{p^2}e^{ip(x-y)} = \int \frac{d^4p}{(2\pi)^4}\,\frac{+\not p}{p^2}e^{ip(x-y)}$$

change $p \to -p$ (propagator from x to y)

$$-S(x-y) = \int \frac{d^4p}{(2\pi)^4} \frac{-\not{p}}{p^2} e^{-ip(x-y)} = \int \frac{d^4p}{(2\pi)^4} \frac{-\not{p}}{p^2} e^{ip(y-x)} .$$

So it is equivalent to invert the propagator $y \to x$ to $x \to y$ with positive energy and negative mass. Therefore it reduces the charges, currents, energies, ... of the particles and antiparticles, we have

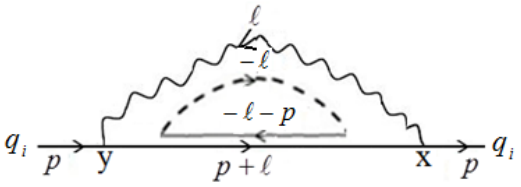

Fig. 5: Omitting the distance x-y from the propagator.

$$(p+\ell) + (-p-\ell) = 0 \text{ and } (-\ell) + (\ell) = 0$$

so incoming with p and outcoming with p, it is like to say the particles jump from y to x, in other words the distance $y - x$ is removed from the interaction. We expect the fields dual behavior takes place in negative potential. If there is no negative potential the paired particles would not survive (never condense). For the quarks, the case $0 < r < a$ must associate with negative potential u and $E + u < 0$. Because the behavior of the strong interaction coupling constant at low energy α_s is high, we expect negative potential at low energy $E + u < 0$ $(E > 0, u < 0)$, so the quarks condense.

Because of the dual behavior of the quarks field which leads to quarks composite in scalar charged particles like the pions (π^-, π^0, π^+) and because of their quantized charges $(-1, 0, +1)$ we expect the hadrons charges to be also quantized $(-Q, -Q+1, \ldots 0, +1, \ldots, +Q)$ this quantization relates to the dual behavior of the quarks field in different hadrons, pairing quarks of different hadrons, so these condensed quarks; pions, kaons, ... are shared between the hadrons, so we put them together with the hadrons in groups, like the pions $(-1, 0, +1)$ which can be inserted in $SU(2)$ generators which can represent the proton-neutron pairing. Therefore the protons and neutrons Lagrangian contains the terms $-ig_{\pi N}\pi^\alpha \bar{N} T_2^\alpha N$ with the nucleon field $N = \begin{pmatrix} p \\ n \end{pmatrix}$.

3 The quarks plasma

We tried before to explain how the quarks are confined, for the strong interaction, we have the condition $r < a \neq 0$ at low limited energy and the condition $r > a \to 0$ at high energies for free quarks where the length a is removed from the propagators. But it appears to be fixed at low limited energy. In the last section we showed there is dual behavior for the quarks field, but when the length a is fixed, the result is scalar particles (pions) with mass $1/a_0$ at low limited energy and the result is the chiral symmetry breaking. We found the length a appears in the quark-quark strong interaction (gluons exchanging) potential $U(r)_{r<a} < 0$, so it relates to interaction strength. That is because the behavior of the length a is like the behavior of the coupling constant α_s. The confinement (at low energy $r < a$) means when $r \to a$ the two interacting quarks kinetic energy becomes zero (ignore the quark mass), therefore the highest kinetic energy the quark can get equals σa which relates to the potential $U(r) = -u_0 + \sigma r + \ldots$ for $r < a$ (at low limited energy). When $r \to a$ the potential becomes $U(0) = -u_0 = -4\alpha_s/3a < 0$ therefore the total quarks energy becomes negative. In this situation the free quarks disappear ($\mu_0 \to 0$), they become condensed in the hadrons.

We try here to use statistical thermodynamics to show how the free quarks disappear at low energies (low temperatures) where the length a becomes fixed, so the chiral symmetry breaking and the quarks condensation. One of the results is that the confinement phase (3.14) not necessarily associates with chiral symmetry breaking, that is, the chiral symmetry breaking appears at the end of the cooling process when the expanding and cooling are ended and the length a becomes fixed, therefore the chiral symmetry breaking occurs and the pions become massive $m = 1/a_0$.

We start with the massless quarks, their energy in volume V is

$$E = c \int_{a^3} d^3r \int_0^\infty d\varepsilon\, g(\varepsilon)\, \varepsilon \frac{1}{e^{\beta(\varepsilon-\mu(r))}+1} \quad : g(\varepsilon) = g_q \frac{V}{2\pi^2} \varepsilon^2 \tag{3.1}$$

where $\mu(r) = \mu_0 + u(r)$ with $u(r) = -\frac{4\alpha_s}{3r}\left(1 - e^{-r/a}\right)$.

Here we inserted the quark-quark strong interaction potential $U(r)$ in the chemical potential (for decreasing the free quarks energy, as we think, the quarks potential reduces the free quarks chemical potential and make them condense at low energy) and because $r < a$ we integrate over the volume a^3: r is the distance between the interacting quarks. We can replace $4\alpha_s/3 \to \alpha_s$.

The constant c is determined by comparing with free quarks high energy where the potential $U(r) \to 0$ and $\alpha_s \to 0$ (decoupling) at high energies, so the length $a \to 0$ that is as we said before, the behavior of the length a is like the behavior of the coupling constant g_s therefore the quarks become free at high energies.

By integrating over the energy (Maple program) we have:

$$\begin{aligned} E &= c g_q \frac{V}{2\pi^2} \int_{a^3} d^3r \int_0^\infty d\varepsilon \frac{\varepsilon^3}{e^{\beta(\varepsilon-\mu(r))}+1} \\ &= c g_q \frac{V}{2\pi^2\beta^4} \int_{a^3} d^3r \left[\frac{7\pi^4}{60} + \frac{\pi^2}{2} u_0(r)^2 + \frac{1}{4} u_0(r)^4 + \right. \\ &\qquad \left. + 6 \sum_{k=1}^{\infty} \frac{(-1)^k e^{-k\beta\mu(r)}}{k^4} \right] \end{aligned}$$

with $u_0(r) = \beta\mu(r) = \beta(\mu_0 + u(r))$. By integrating over r (the distance between the interacting quarks) we have

$$E = cg_q \frac{2Va^3}{\pi x^4}\Bigg[3.78 + 2\,(\beta\mu_0)^2\Bigg(0.82 - 1.16\,\frac{\alpha_s}{a\mu_0} + \\ + 0.41\left(\frac{\alpha_s}{a\mu_0}\right)^2\Bigg) + (\beta\mu_0)^4\Bigg(0.08 - 0.23\,\frac{\alpha_s}{a\mu_0} + \\ + 0.25\left(\frac{\alpha_s}{a\mu_0}\right)^2 - 0.12\left(\frac{\alpha_s}{a\mu_0}\right)^3 + 0.02\left(\frac{\alpha_s}{a\mu_0}\right)^4\Bigg) + \\ + 6\sum_{k=1}^{\infty}\int_0^1 x^2dx\,\frac{(-1)^k e^{-k\beta\mu(x)}}{k^4}\Bigg].$$

g_q is the quarks degeneracy number and $x = \beta\mu_0$. For easiness we write $\alpha_s/a\mu_0 = 2\sigma a/\mu_0 = y$ in the energy relation. So it becomes

$$E = cg_q \frac{2Va^3}{\pi x^4}\Big[3.78 + 2(\beta\mu_0)^2\left(0.82 - 1.16y + 0.41y^2\right) \\ + (\beta\mu_0)^4\left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right) \\ + 6\sum_{k=1}^{\infty}\int_0^1 x^2dx\,\frac{(-1)^k e^{-k\beta\mu(x)}}{k^4}\Big]. \qquad (3.2)$$

at high energy: $x = \beta\mu_0 = \mu_0/T \to 0$. To find the constant c we compare with quarks high energy where they are free massless particles:

$$E_{high} = g_q V\,\frac{7\pi^2}{240}\,T^4\,.$$

When T is high, $x = (\mu_0/T) \to 0$ and $y \to 0$ therefore $\beta\mu(x) \to 0$ so we expand $e^{-k\beta\mu(x)}$ near $\beta\mu(x) = 0$, we have:

$$E_{high} = cg_q\frac{2a^3V}{\pi x^4}\,[3.78 - 1.88 + O(x,y)] \to cg_q\frac{2a^3V}{\pi x^4}1.9$$

$$\to g_q\frac{7\pi^2V}{240}T^4 = cg_q\frac{2a^3V}{\pi x^4}1.9 \to c = \frac{\pi}{2a^3 1.9}\frac{7\pi^2}{240}\mu_0^4 \qquad (3.3)$$

The energy becomes:

$$E = \frac{1}{1.9}\,\frac{7\pi^2}{240}\,\mu_0^4 g_q\,\frac{V}{(\beta\mu_0)^4}\Big[3.78 + 2(\beta\mu_0)^2(0.82 - 1.16y + \\ + 0.41y^2) + (\beta\mu_0)^4(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + \\ + 0.02y^4) + 6\sum_{k=1}^{\infty}\int_0^1 x^2dx\,\frac{(-1)^k e^{-ku_0(x)}}{k^4}\Big].$$

Now we see the effects of the length a on the energy, at high energy, by fixing $x = \mu_0/T$ and varying $y = \sigma a/2\mu_0 < 1$:

$$E_{high} = \frac{1}{1.9}\frac{7\pi^2}{240}\,g_q V\mu_0^4\,x^{-4}\Big[1.9 + x(1.8 - 1.24y) + \\ + x^2(0.82 - 1.18y + 0.42y^2) + x^3(0.23 - 0.47y + \\ + 0.33y^2 - 0.08y^3) + x^4(0.04 - 0.12y + \\ + 0.13y^2 - 0.07y^3 + 0.01y^4) + \ldots\Big]_{x=\beta\mu_0\to 0}\,. \qquad (3.4)$$

We expanded $e^{-k\beta\mu(x)}$ near $\beta\mu(x) = 0$ and fixed the tension σ as we assumed before, so we have Fig. 6.

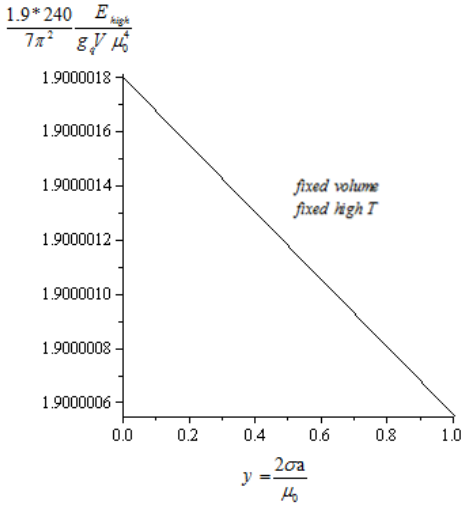

Fig. 6: Decreasing the high energy with increasing y.

It appears in the figure that the high energy quarks lose energy when the length a increases although the temperature is fixed. That means, when the length a increases the number of the excited quarks decreases. That is because of the attractive linear potential $\sigma r\ldots$ between the quarks, that potential absorbs an energy ($r < a$ confinement, section 1), so the quarks are cooled faster by the expansion. As we said before, the behavior of length a is like the behavior of the coupling constant α_s so when the energy dropped to lowest energy, the length a increased extremely and this is fast cooling (extreme cooling). That occurs when the particles spread away, the length a, as a distance between the quarks, increases.

To determine the end, we search for the balance situations, such as zero pressure, confinement condition, ... First we find the high energy pressure including the effects of the potential σa. Starting from the general pressure relation:

$$p = -\frac{\partial}{\partial V}F \text{ where } F = -T\ln Z = -\frac{1}{\beta}\,\ln Z$$

here we use the relation:

$$\ln Z = c\int_{a^3} d^3r\int_0^{\infty} d\varepsilon\, g(\varepsilon)\ln\left(e^{-\beta(\varepsilon-\mu(r))}+1\right)$$

$$: \; g(\varepsilon) = g_q\,\frac{V}{2\pi^2}\,\varepsilon^2$$

and the pressure becomes

$$P = \frac{1}{3}\,\frac{\partial}{\partial V}E$$

so for high energy $x = \beta\mu_0 \to 0$ we have the pressure:

$$\begin{aligned} P_{high} &= \frac{1}{3}\frac{\partial}{\partial V}E_{high} \\ &= \frac{\partial}{\partial V}\frac{1}{3\times 1.9}\frac{7\pi^2}{240}\,g_q V\mu_0^4\,x^{-4}\Big[1.9 + x(1.8 - 1.24y) + \\ &+ x^2(0.82 - 1.18y + 0.42y^2) + x^3(0.23 - 0.47y + \\ &+ 0.33y^2 - 0.08y^3) + x^4(0.04 - 0.12y + \\ &+ 0.13y^2 - 0.07y^3 + 0.01y^4) + \dots\Big]. \end{aligned}$$

Now the key point is, we want to include the potential effect on the pressure so we replace the volume V with the volume $a^3 \sim y^3$ so

$$\begin{aligned} P_{high} \to \frac{\partial}{\partial y^3}y^3\frac{1}{3\times 1.9}\frac{7\pi^2}{240}\,g_q\,\mu_0^4\,x^{-4}\Big[1.9 + x(1.8 - 1.24y) + \\ + x^2(0.82 - 1.18y + 0.42y^2) + x^3(0.23 - 0.47y + \\ + 0.33y^2 - 0.08y^3) + x^4(0.04 - 0.12y + \\ + 0.13y^2 - 0.07y^3 + 0.01y^4) + \dots\Big] \end{aligned} \tag{3.5}$$

which is represented in Fig. 7, without conditions on y or on the length a.

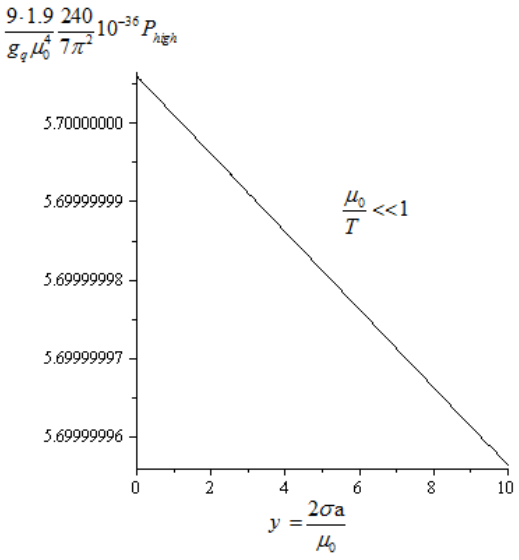

Fig. 7: The effects of potential σa on the pressure.

It is clear (without conditions on y) the pressure decreases with increasing the length a (decreasing the quarks energy $-p^2$) until it becomes zero, then negative. That becomes clear at low energy where there are conditions on y and so on the length a.

For the low energy quarks, $T \to 0$ so $\beta\mu(x) \to \infty$ so $e^{-k\beta\mu(x)} \to 0$. The energy becomes:

$$\begin{aligned} E_{low} = \frac{1}{1.9}\frac{7\pi^2}{240}\,\mu_0^4\,g_q\,\frac{V}{(\beta\mu_0)^4}\Big[3.78 + \\ + 2\,(\beta\mu_0)^2\left(0.82 - 1.16y + 0.41y^2\right) + \\ + (\beta\mu_0)^4\left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right)\Big]. \end{aligned} \tag{3.6}$$

Making $x = T/\mu_0$ so

$$\begin{aligned} E_{low} = \frac{1}{1.9}\frac{7\pi^2}{240}\,\mu_0^4\,g_q V x^4\Big[3.78 + 2x^{-2}\left(0.82 - 1.16y + 0.41y^2\right) \\ + x^{-4}\left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right)\Big]. \end{aligned}$$

Now the key point, we want to show the effect of the potential σa on the energy so we see the behavior of the energy in the volume a^3 with respect to $y = 2\sigma a/\mu_0$ the diagram is given in Fig. 8. That is extreme behavior after $y = 0.6$ where the

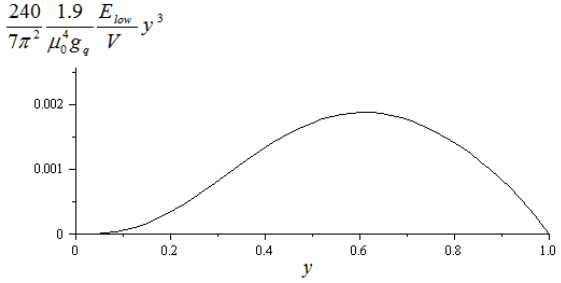

Fig. 8: The extremely decreasing in quarks low energy in the strong interaction.

energy $(E/V)\,a^3$ decreases when the volume a^3 increases, the end in $y = 1$ where the free quarks disappear for $y > 1$.

Now we can distinguish between the confinement and the chiral symmetry breaking, when $y > 0.6$ there is confinement: extreme cooling, negative pressure. But when reach $y = 1$ there is chiral symmetry breaking where the length a becomes fixed, and from the quarks field dual behavior there are scalar charged particles with mass $1/a$ appear when the length a is fixed with non-zero value a_0. Here the evidence for fixing the length a is the lowest limited quarks energy, that is as we said before, the behavior of the length a is like the behavior of the coupling constant α_s so when the quarks energy dropped (extreme cooling) the length a increases extremely to reach the highest value when $y = 1$ which is equivalent to smallest energy $E = 0$ (the cooling end). Another evidence for fixing the length a (chiral symmetry breaking) is the low energy pressure:

$$P_{low} = \frac{1}{3}\frac{\partial}{\partial V}E_{low} \to \frac{1}{3}\frac{\partial}{\partial y^3}\frac{E_{low}}{V}\,y^3\,.$$

To include the potential effect we study the pressure using the

volume $a^3 \sim y^3$ therefore

$$P_{low} \to \frac{1}{3}\frac{\partial}{\partial y^3}\frac{1}{1.9}\frac{7\pi^2}{240} g_q \mu_0^4 y^3 x^4 \left[3.78 + \right.$$
$$+ 2x^{-2}\left(0.82 - 1.16y + 0.41y^2\right)$$
$$\left. + x^{-4}\left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right)\right]$$

and therefore

$$\frac{P_{low}}{\mu_0^4} = \frac{1}{9\times 1.9}\frac{7\pi^2}{240} g_q \left[3\times 3.78\, x^4 + \right.$$
$$+ 3\times 2\times x^2\left(0.82 - 1.16y + 0.41y^2\right) \qquad (3.7)$$
$$+ 3\times(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4) +$$
$$+ 2yx^2(-1.16 + 0.82y) +$$
$$\left. + y\left(-0.23 + 0.5y - 0.36y^2 + 0.08y^3\right)\right].$$

We see its behavior in Fig. 9 below

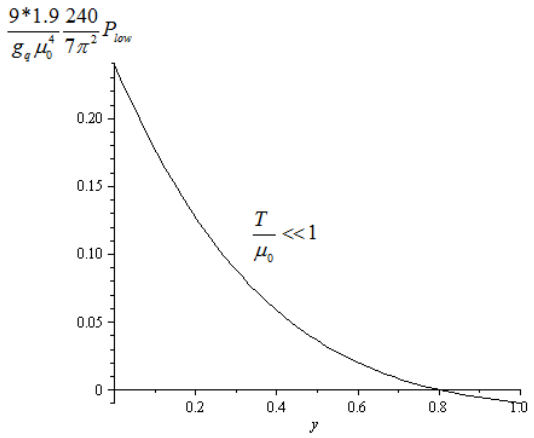

Fig. 9: The extremely decreasing in the pressure at low energy.

It is clear from the figure, when $y > 0.6$ the quarks pressure becomes negative. We expect the condensed quarks phase (confinement quarks) has positive pressure, so the preferred phase is the condensed quarks phase. So when $y > 0.6$ the quarks condense until $y = 1 : a \to a_0 \approx 1/(135 - 140\,\text{Mev})$ the quarks disappear, the scalar charged particles (pions) appear instead of them, that is because of the quarks dual behavior (free-condensed quarks), but at low limited energy the condensed phase has a big chance instead the free phase.

3.1 The confinement phase

In this paper we study two quarks (up and down) condensation in the pions (π^0, π^+, π^-) and baryons (n, p^+, p^-), so the degeneracy number is $g_q = 2_{flavor}\times 2_{charge}\times 2_{spin}\times 3_{color} = 24$.

We need more clarifying for determining if the quarks could stay free particles or they condense in hadrons. We can think they could be free if their energy is enough for covering the strong interaction potential and stay free particles with least possible energy (at 0 temperature). Unless they condense in the hadrons.

To cover the strong interaction potential means to lose an energy E_u which is transferred to the exchanged static gluons and pions which are created between the low energy quarks. So the remaining energy in the volume $4\pi a^3/3$ is

$$\frac{E_{q,low}}{V}\frac{4\pi}{3}a^3 - \frac{E_u}{V}\frac{4\pi}{3}a^3 . \qquad (3.8)$$

This energy must be enough for the least possible free quarks. Therefore we must determine the chemical potential μ_0 of the free quarks with smallest possible density at 0 temperature.

According to the quarks confinement $r < a$ at low limited energy, which means the highest possible distance between the two interacting quarks is a, we expect the least quarks density is two quarks in the volume $4\pi\,(a/2)^3/3$.

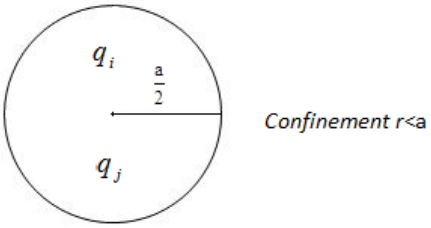

Fig. 10: The quarks confinement at low energy.

From this view we can calculate the least quarks chemical potential μ_0 of free quarks:

$$2\left(\frac{4\pi}{3}\left(\frac{a}{2}\right)^3\right)^{-1} = \frac{1}{V}\int_0^{\mu_0} g(\varepsilon)\,d\varepsilon = g_q\frac{\mu_0^3}{6\pi^2}$$
$$\to \left(\mu_0\frac{a}{2}\right)^3 = \frac{9\pi}{g_q}$$
$$\to (\mu_0 a)^3 = \frac{8\times 9\pi}{g_q}$$

$1/a$ is the pion mass when $a \to a_0$ in the end of free quarks phase so $1/a \to (135 - 140)\,\text{Mev}$. So the least free quarks energy density in 0 temperature is

$$\frac{\varepsilon_{free}}{V} = \frac{1}{V}\int_0^{\mu_0} g(\varepsilon)\,\varepsilon\,d\varepsilon = g_q\frac{\mu_0^4}{4\times 2\pi^2} .$$

The smallest energy of the free quarks in the volume $4\pi a^3/3$ is

$$\varepsilon_{free,a^3} = g_q\frac{\mu_0^4}{4\times 2\pi^2}\frac{4\pi a^3}{3} = \frac{4\pi}{3}g_q\frac{\mu_0}{4\times 2\pi^2}(\mu_0 a)^3$$
$$= \frac{4\pi}{3}g_q\frac{\mu_0}{4\times 2\pi^2}\frac{8\times 9\pi}{g_q} = \frac{4\pi}{3}\frac{9}{\pi}\mu_0 \qquad (3.9)$$

therefore

$$\frac{\varepsilon_{free,a^3}}{2\mu_0} = \frac{4\pi}{3}\frac{9}{2\pi} = \frac{4\pi}{3}\times 1.43\,.$$

Because the chemical potential $\mu \sim 1/a$ and $\mu \to \mu_0$ when $a \to a_0$ and because $y \sim a$ so we modified $\mu_0 \to \mu_0/y$ so

$$\frac{4\pi}{3}\times 1.43 \to \frac{4\pi}{3}\times\frac{1.43}{y}\,. \tag{3.10}$$

Now we find the least energy E_u which is transferred to the static exchanged gluons and pions according to the potential

$$u(r) = -\frac{4\alpha_s}{3r}\left(1-e^{-r/a}\right) \approx -u_0 + \sigma r : r < a\,.$$

We absorbed $4/3$ to α_s so and made $\alpha_s/a\mu_0 = 2\sigma a/\mu_0 = y$ the constant σ is the string tension. This potential is inserted to reduce the chemical potential μ_0 and the energy is renormalized at high energy. So we have $\mu_0 \to \mu_0 + u(r)$:

$$\mu(r) = \mu_0 - \frac{\alpha_s}{r}\left(1-e^{-r/a}\right) \approx \mu_0 - u_0 + \sigma r : r < a\,.$$

Therefore we can calculate the least absorbed energy by this potential, by calculating the changes on the energy density at 0 temperature

$$\begin{aligned}\frac{\varepsilon(\alpha_s/a)}{V} = \frac{\varepsilon(y)}{V} &= \frac{1}{V}\,c\int_0^a 4\pi r^2 dr\int_0^{\mu(r)} g(\varepsilon)\,\varepsilon\,d\varepsilon\\ &= c\int_0^a 4\pi r^2 dr\, g_q\,\frac{\mu(r)^4}{4\times 2\pi^2}\,.\end{aligned}$$

The constant c is determined

$$c = \frac{\pi}{2a^3 1.9}\,\frac{7\pi^2}{240}$$

so the interaction energy is

$$\begin{aligned}\frac{\varepsilon(\alpha_s/a)}{V} = \frac{\varepsilon(y)}{V} &= g_q\,\frac{\pi}{2a^3 1.9}\,\frac{7\pi^2}{240}\,\frac{4\pi}{4\times 2\pi^2}\int_0^a r^2 dr\mu(r)^4\\ &= g_q\,\frac{7\pi^2}{4\times 1.9\times 240a^3}\int_0^a r^2 dr\mu(r)^4\,.\end{aligned}$$

This becomes

$$\begin{aligned}\frac{\varepsilon(y)}{V} &= g_q\,\frac{7\pi^2}{4\times 1.9\times 240a^3}\int_0^a r^2 dr\left[\mu_0 - \frac{\alpha_s}{r}\left(1-e^{-r/a}\right)\right]^4\\ &= g_q\,\frac{7\pi^2}{4\times 1.9\times 240a^3}(\mu_0)^4\int_0^a r^2 dr\left[1-\frac{\alpha_s}{\mu_0 r}\left(1-e^{-r/a}\right)\right]^4\end{aligned}$$

Using the change $r = ax$ so

$$\begin{aligned}\frac{\varepsilon(y)}{V} =& g_q\,\frac{7\pi^2}{4\times 1.9\times 240a^3}\,(\mu_0)^4\\ &\int_0^1 a^3x^2dx\left[1-\frac{\alpha_s}{\mu_0 ax}\,(1-e^{-x})\right]^4\end{aligned}$$

therefore

$$\frac{\varepsilon(y)}{V} = g_q\,\frac{7\pi^2}{4\times 1.9\times 240}\,(\mu_0)^4\int_0^1 x^2dx\left[1-\frac{y}{x}\,(1-e^{-x})\right]^4\,.$$

The spent energy for the interaction in the volume $4\pi a^3/3$ is

$$\begin{aligned}\varepsilon_{u,a^3} =& \frac{\varepsilon(1)-\varepsilon(0)}{V}\,\frac{4\pi a^3}{3}\\ =&\frac{4\pi}{3}\,g_q\,\frac{7\pi^2(\mu_0 a)^3\mu_0}{4\times 1.9\times 240}\\ &\left(\int_0^1 x^2dx\left[1-\frac{1}{x}\,(1-e^{-x})\right]^4 - \int_0^1 x^2dx\right)\end{aligned} \tag{3.11}$$

and it becomes

$$\begin{aligned}\varepsilon_{u,a^3} &= -\frac{4\pi}{3}\,g_q\,\frac{7\pi^2}{4\times 1.9\times 240}\,(\mu_0 a)^3\mu_0\,0.33\\ &= -\frac{4\pi}{3}\,g_q\,\frac{7\pi^2}{4\times 1.9\times 240}\,\frac{8\times 9\pi}{g_q}\,\mu_0\times 0.33\,.\end{aligned}$$

Therefore

$$\begin{aligned}\varepsilon_{u,a^3} &= -\frac{4\pi}{3}\,g_q\,\frac{7\pi^2}{4\times 1.9\times 240}\,(\mu_0 a)^3\mu_0\times 0.33\\ &= -\frac{4\pi}{3}\,\frac{7\times 8\times 9\times 0.33\pi^3}{4\times 1.9\times 240}\,\mu_0 = -\frac{4\pi}{3}\times 2.82\,\mu_0\,.\end{aligned} \tag{3.12}$$

So we have

$$\frac{\varepsilon_{u,a^3}}{2\mu_0} = -\frac{4\pi}{3}\times 1.41\,.$$

As for E_{free} we replace

$$\frac{4\pi}{3}\times 1.41 \to \frac{4\pi}{3}\times\frac{1.41}{y}\,.$$

Now we find the confinement condition at any temperature, if the quarks energy is not enough to cover the interaction energy E_u and give free quarks with smallest density, at 0 temperature, then they become confinement ($r < a$), so the confinement condition

$$E(T,y) - \varepsilon_u - \varepsilon_{free} < 0\,. \tag{3.13}$$

Then

$$\frac{E(T,y)}{V}\,\frac{4\pi a^3}{3} - \frac{\varepsilon_u}{V}\,\frac{4\pi a^3}{3} - \frac{\varepsilon_{free}}{V}\,\frac{4\pi a^3}{3} < 0$$

or

$$\frac{E(T,y)}{2\mu_0 V}\,\frac{4\pi a^3}{3} - \frac{\varepsilon_u}{2\mu_0 V}\,\frac{4\pi a^3}{3} - \frac{\varepsilon_{free}}{2\mu_0 V}\,\frac{4\pi a^3}{3} < 0\,.$$

We consider

$$\sigma_{a^3}a = \frac{\varepsilon_{free}}{V}\,\frac{4\pi a^3}{3}$$

as critical energy of free quarks for lowest energy, the tension σ_{a^3} here is the volume tension. Therefore this critical energy

is transferred to the produced hadrons and photons. Using the quarks low energy

$$E_{low} = \frac{1}{1.9}\frac{7\pi^2}{240} g_q \mu_0^4 V x^4 \left[3.78 + \right.$$
$$+ 2\, x^{-2}\left(0.82 - 1.16y + 0.41y^2\right) +$$
$$\left. + x^{-4}\left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right)\right].$$

With $x = T/\mu_0 \ll 1$, the confinement condition becomes

$$\frac{1}{2}\frac{1}{1.9}\frac{7\pi^2}{240}\mu_0^3 g_q\, x^4\, \frac{4\pi a^3}{3}\left[3.78 + \right.$$
$$+ 2x^{-2}\left(0.82 - 1.16y + 0.41y^2\right) +$$
$$\left. + x^{-4}\left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right)\right] -$$
$$- \frac{\varepsilon_u}{2\mu_0 V}\frac{4\pi a^3}{3} - \frac{\varepsilon_{free}}{2\mu_0 V}\frac{4\pi a^3}{3} < 0\,.$$

It becomes

$$\frac{1}{2}\frac{1}{1.9}\frac{7\pi^2}{240} g_q \frac{4\pi(\mu_0 a)^3}{3}\left[3.78x^4 + \right.$$
$$+ 2x^2\left(0.82 - 1.16y + 0.41y^2\right) +$$
$$\left. + \left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right)\right] -$$
$$- \frac{4\pi}{3}\times\frac{1.41}{y} - \frac{4\pi}{3}\times\frac{1.43}{y} < 0\,.$$

We had the relation

$$(\mu_0 a)^3 = \frac{8\times 9\pi}{g_q}\,,$$

therefore, the condition becomes

$$\frac{1}{2}\frac{1}{1.9}\frac{7\pi^2}{240}\frac{4\times 8\times 9\pi^2}{3}\left[3.78x^4 + \right.$$
$$+ 2\,x^2\left(0.82 - 1.16y + 0.41y^2\right)$$
$$\left. + \left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right)\right] -$$
$$- \frac{4\pi}{3}\times\frac{1.41}{y} - \frac{4\pi}{3}\times\frac{1.43}{y} < 0\,.$$

It becomes

$$3.78x^4 + 2x^2\left(0.82 - 1.16y + 0.41y^2\right) +$$
$$+ \left(0.08 - 0.23y + 0.25y^2 - 0.12y^3 + 0.02y^4\right) - \quad (3.14)$$
$$- 0.16y^{-1} < 0$$

with the curve of Fig. 11.

The critical situation x_c with $y \to 1$ (the end of the extreme cooling)

$$3.78x_c^4 + 2\times 0.07x_c^2 - 0.16574 = 0 \to x_c = 0.438\,.$$

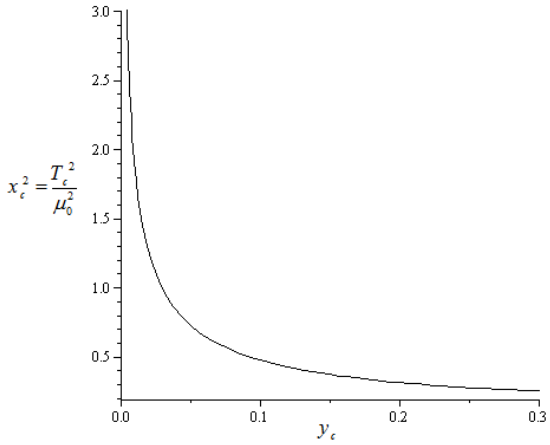

Fig. 11: The critical $x_c^2 y_c$ curve separates the free and confinement quarks phases.

So the critical temperature of the confinement condition when $y \to 1$ from $x_c = T_c/\mu_0$ is $T_c = 0.438\mu_0$. We determine μ_0 from

$$(\mu_0 a)^3 = \frac{8\times 9\pi}{g_q}$$

when $y \to 1$ so $a \to a_0$ we set $1/a_0$ = pion mass = (135 − 140) Mev so therefore, the condition becomes

$$\mu_0 = \frac{1}{a}\left(\frac{8\times 9\pi}{g_q}\right)^{\frac{1}{3}} \to 135\left(\frac{8\times 9\pi}{24}\right)^{\frac{1}{3}}$$
$$= 285.15\,\text{Mev} \quad \text{for}\ \frac{1}{a_0} = 135\,\text{Mev}. \quad (3.15)$$

Hence the critical temperature is $T_c = 0.438 \times 285.15 = 124.9$ Mev.

Now we try to find the produced hadrons, after covering the potential (3.12), the quarks critical energy (possible smallest energy) E_{free} (3.9) is transferred to the produced hadrons and photons. The key idea here is: because the cooling is an extreme cooling, it is expanding $a : 0 \to a_0 = 1/(135 - 140\,\text{Mev})$ so this process is thermally isolated from the other fields (adiabatic change), therefore the produced particles are in $T_c = 124.9$ Mev. We assume that the produced particles are hadrons (fermions and bosons) and photons. When $a : 0 \to a_0 : y \to 1$ the pions become massive $m = 1/a_0$ so we expect the other hadrons become massive at this stage, we assume that is in $T \to T_c$.

Therefore we assume when $T > T_c$ massless hadrons and $T < T_c$ massive hadrons. Anyway in $x_c y_c$ curve we find the confinement is possible at high energy ($T \gg T_c : a \to 0$). First we write using (3.9)

$$\frac{\varepsilon_{free}}{V} = \frac{\varepsilon_{free,a^3}}{4\pi a^3/3} = g_q\frac{\mu_0^4}{4\cdot 2\pi^2} \quad (3.16)$$
$$= \frac{\sigma_{a^3}\, a}{4\pi a^3/3} \to \frac{E_{hadrons} + E_{photons}}{V}\ \text{below}\ x_c y_c\ \text{curve}$$

or

$$\frac{\sigma_{a^3}a}{4\pi a^3/3} = g_q\frac{\mu_0^4}{4\cdot 2\pi^2} \to \varepsilon_f + \varepsilon_b + \varepsilon_{ph}\,.$$

With the densities

$$\varepsilon_f = \frac{E_f}{V}\,,\ \varepsilon_b = \frac{E_b}{V} \text{ and } \varepsilon_{ph} = \frac{E_{ph}}{V}$$

for spin 1/2 hadrons (fermions p^+, p^-, n), spin 0 hadrons (bosons π^0, π^-, π^+) and photons densities. For massless phase $T \gg T_c$ and $y_c \approx 0$ ignoring the chemical potential we have

$$\begin{aligned} n_f &= \frac{N_f}{V} = g_f\,\frac{3\zeta(3)}{4\pi^2}\,T^3\,,\\ n_b &= \frac{N_b}{V} = g_b\,\frac{\zeta(3)}{\pi^2}\,T^3 \text{ and}\\ \varepsilon_{ph} &= \frac{E_{ph}}{V} = g_{ph}\,\frac{\pi^2}{30}\,T^4\,. \end{aligned} \tag{3.17}$$

Now the key point, because the cooling is extreme cooling, to take all the particles (quarks) from high temperature and put them at low temperature, so the same structure at high energy will be at low energies, like the charges ratios, energy distribution over the particles, spins, ... At $T \to T_c$ and $y_c = 1$ the hadrons become massive, we approximate: for bosons (pions with mass $1/a_0 = 135 - 140$ Mev) the energy density becomes:

$$\begin{aligned} &\varepsilon_b = g_b\,\frac{\pi^2}{30}\,T^4 \to \varepsilon_b = g_b\,\frac{\pi^2}{30}\,T^4 + m_{pion}n_b\\ &\text{with } n_b = \frac{N_b}{V} = g_b\,\frac{\zeta(3)}{\pi^2}\,T^3 \text{ and } m_{pion} = \frac{1}{a_0}\,. \end{aligned}$$

And for fermions (let them be p^+, p^-, n) we approximate (ignoring the chemical potential)

$$\begin{aligned} &\varepsilon_f = g_f\,\frac{7}{8}\,\frac{\pi^2}{30}\,T^4 \to \varepsilon_f = g_f\,\frac{7}{8}\,\frac{\pi^2}{30}\,T^4 + m_f n_f\\ &\text{with } n_f = \frac{N_f}{V} = g_f\,\frac{3\zeta(3)}{4\pi^2}\,T^3\,. \end{aligned}$$

So (3.15) becomes

$$\begin{aligned} &\frac{\sigma_{a^3}a}{4\pi a^3/3} = g_q\,\frac{\mu_0^4}{4\cdot 2\pi^2} = \varepsilon_f + \varepsilon_b + \varepsilon_{ph}\\ &= g_f\,\frac{7}{8}\,\frac{\pi^2}{30}T_c^4 + m_f n_f + g_b\,\frac{\pi^2}{30}T_c^4 + \frac{1}{a_0}n_b + g_{ph}\,\frac{\pi^2}{30}T_c^4\,. \end{aligned} \tag{3.18}$$

with $g_{quarks} = 2_{flavor}\times 2_{charge}\times 2_{spin}\times 3_{color}$, $g_f = 3_{charge}\times 2_{spin}$, $g_b = 3_{charge}$ and $g_{ph} = 2_{polarization}$.

Now we calculate (3.17) for $1/a_0 = 135\,\text{Mev}\,(\pi^0)$, $\mu_0 = 285.15$ Mev, and $T_c = 124.9$ Mev we have

$$\begin{aligned} &2.0096\times 10^9\,\text{Mev}^4 = 6\times\frac{7}{8}\,\frac{\pi^2}{30}(124.9)^4 +\\ &+\ m_f 6\times\frac{3\zeta(3)}{4\pi^2}(124.9)^3 + 3\times\frac{\pi^2}{30}(124.9)^4 +\\ &+ 135\times 3\times\frac{\zeta(3)}{\pi^2}(124.9)^3 + 2\times\frac{\pi^2}{30}(124.9)^4\,. \end{aligned}$$

Its solution is $m_f = 1023$ Mev. We keep 2.0096×10^9 Mev^4 as smallest possible energy density.

For $1/a_0 = 140$ Mev, $\mu_0 = 295.7$ Mev so $T_c = 129.5$ Mev the mass m_f becomes $m_f = 798.4$ Mev. Therefore it must be $135\,\text{Mev} < 1/a_0 < 140\,\text{Mev}$.

For $1/a_0 = 136.8$ Mev we have $T_c = 126.56$ Mev then the mass m_f becomes $m_f \approx 938$ Mev so the fermions (hadrons) are the baryons (p^+, p^-, n).

Therefore we fix it $1/a_0 = 136.8$ Mev, we use it to cancel the dark matter. Maybe there is an external pressure $-P_{ex}$ so the lost energy is $P_{ex}\,4\pi a^3/3$.

Now we try to calculate the ratio N_q/N_h. From the condensation relation

$$N_q\delta\mu_q + N_h\delta\mu_h = 0$$

N_h is the hadrons (consider only the fermions) and μ_h is their chemical potential.

We assumed before the relation for the quarks chemical potential

$$\mu(r) = \mu_0 + u(r)\ \text{ with } u(r) = -\frac{\alpha_s}{r}\left(1 - e^{-r/a}\right)$$

$$\text{so } \delta\mu_q(r) = u(r) = -\frac{\alpha_s}{r}\left(1 - e^{-r/a}\right)\,.$$

The effect of this changing appeared in $y = \alpha_s/a\mu_0$ in the results. For the hadrons we have

$$\delta\mu_h = -\frac{N_q}{N_h}\,\delta\mu_q = -\frac{N_q}{N_h}\,u(r)\,.$$

That is right if we consider the hadrons are massless, that is when $T \gg T_c$ and $y \ll 1$ (in the condensation phase, below the curve $x_c y_c$) so we have the chemical potential for the hadrons

$$\mu_h(r) = \mu_{0h} - u(r)\ \text{ with } u(r) = -\frac{\alpha_s}{r}\left(1 - e^{-r/a}\right)$$

therefore we replace $y \to (-N_q\mu_{0q}/N_h\mu_{0h})\,y$ in the quarks energy to get the hadrons energy. The energy of the hadrons becomes

$$\begin{aligned} &E_{H,low} = \frac{1}{1.9}\,\frac{7\pi^2}{240}\,\mu_{0h}^4\,g_h\,V\,x^4\Bigg[3.78 +\\ &+ 2x^{-2}\left(0.82 + 1.16\left(\frac{N_q\mu_{0q}}{N_h\mu_{0h}}\right)y + 0.41\left(\frac{N_q\mu_{0q}}{N_h\mu_{0h}}\right)^2 y^2\right) +\\ &+ x^{-4}\left(0.08 + 0.23\left(\frac{N_q\mu_{0q}}{N_h\mu_{0h}}\right)y + 0.25\left(\frac{N_q\mu_{0q}}{N_h\mu_{0h}}\right)^2 y^2 +\right.\\ &\left.+ 0.12\left(\frac{N_q\mu_{0q}}{N_h\mu_{0h}}\right)^3 y^3 + 0.02\left(\frac{N_q\mu_{0q}}{N_h\mu_{0h}}\right)^4 y^4\right)\Bigg]. \end{aligned}$$

Assume $\mu_{0h} = \mu_{0q}$ and $y = 1$ so

$$E_{H,low} = \frac{1}{1.9}\,\frac{7\pi^2}{240}\,\mu_{0q}^4\, g_h\, V\left[3.78x^4 + 2x^2\left(0.82 + \right.\right.$$
$$+1.16\left(\frac{N_q}{N_h}\right) + 0.41\left(\frac{N_q}{N_h}\right)^2\Big) + \left(0.08 + 0.23\left(\frac{N_q}{N_h}\right) + \right.$$
$$\left.\left. + 0.25\left(\frac{N_q}{N_h}\right)^2 + 0.12\left(\frac{N_q}{N_h}\right)^3 + 0.02\left(\frac{N_q}{N_h}\right)^4\Big)\right].$$

So the chemical potential μ_h of the hadrons becomes

$$\mu_h^4 = \mu_{0q}^4\left(1 + \frac{0.23}{0.08}\left(\frac{N_q}{N_h}\right) + \frac{0.25}{0.08}\left(\frac{N_q}{N_h}\right)^2 + \right.$$
$$\left. + \frac{0.12}{0.08}\left(\frac{N_q}{N_h}\right)^3 + \frac{0.02}{0.08}\left(\frac{N_q}{N_h}\right)^4\right).$$

When $T < T_c$ the hadrons become massive, as we assumed before, so for massive hadrons with $m_f = 938$ Mev we expect $\mu_h = m_f = 938$ Mev when they cooled with small densities. Therefore

$$(938)^4 = (285.15)^4\left(1 + \frac{0.23}{0.08}\left(\frac{N_q}{N_h}\right) + \frac{0.25}{0.08}\left(\frac{N_q}{N_h}\right)^2 + \right.$$
$$\left. + \frac{0.12}{0.08}\left(\frac{N_q}{N_h}\right)^3 + \frac{0.02}{0.08}\left(\frac{N_q}{N_h}\right)^4\right).$$

Its positive solution is $N_q/N_h = 3.1$ so they are the baryons (fermions with three quarks). For 0 temperature fermions the chemical potential is approximated by

$$\mu_0^2 = m^2 + \left(\frac{N}{V}\frac{6\pi^2}{g_f}\right)^{2/3}.$$

For low hadrons density we ignored the term

$$\left(\frac{N}{V}\frac{6\pi^2}{g_f}\right)^{2/3}.$$

4 The nuclear compression

The cooled hadrons have high density, so there is hidden high pressure, that pressure makes influence δa so δy near $y = 1$ or it makes $y = 1 + \delta y$: $\delta y \approx 0.005$ so the cooled quarks inside the hadrons fluctuate, this depends on the energy, if the energy is high then there are new hadrons. These processes let the interacting hadrons lose kinetic energy and form the pions.

Because the number of quarks increases although the hadrons are fixed, therefore the hadrons energy decreases and they cannot spread away. We can see how the chemical potential of the interacting hadrons changes under the fluctuation $\delta y \sim \delta a$ (due to the quarks interaction) from the condensation relation $N_q\delta\mu_q + N_h\delta\mu_h = 0$ we have $\delta\mu_h = -N_q\delta\mu_q/N_h$ for the fluctuation δy we have

$$\delta\mu_h = -\frac{N_q}{N_h}\frac{\partial\mu_q}{\partial y}\delta y$$

from quarks chemical potential (4.4), we find

$$\frac{\partial\mu_q}{\partial y} < 0 \quad \text{so} \quad -\frac{\partial\mu_q}{\partial y} > 0$$

therefore we have

$$\delta\mu_h = \frac{N_q}{N_h}\left(-\frac{\partial\mu_q}{\partial y}\right)\delta y < 0 \quad \text{when} \quad \delta y < 0$$

which is the quarks compressing, when the hadrons collide together this leads to $\delta y < 0$ (compression) so the hadrons lose energy and new hadrons are created. And when they try to extend (spread away) $\delta y > 0$ so $\delta\mu_h > 0$, there will be a negative potential.

For the interacting hadrons pressure we have the phase changing relation $V_q\delta P_q + V_h\delta P_h = 0 : V$ volume, we have

$$\delta P_h = -\frac{V_q}{V_h}\,\delta P_q = -\frac{V_q}{V_h}\,\frac{\partial P_q}{\partial y}\,\delta y$$

because $\partial P_q/\partial y < 0 \to -\partial P_q/\partial y > 0$ therefore when the hadrons collide together $\delta y < 0$ so their pressure decreases, they lose energy, so new hadrons are created.

We have

$$\delta y = \left(-\frac{V_q}{V_h}\frac{\partial P_q}{\partial y}\right)^{-1}\delta P_h \quad \text{at} \quad y = 1\,.$$

So the hadrons chemical potential becomes

$$\delta\mu_h = \frac{N_q}{N_h}\left(-\frac{\partial\mu_q}{\partial y}\right)\left(-\frac{V_q}{V_h}\frac{\partial P_q}{\partial y}\right)^{-1}\delta P_h \;:\; y = 1\,.$$

It becomes

$$\delta\mu_h = \frac{N_q V_h}{N_h V_q}\left(\frac{\partial\mu_q}{\partial y}\right)\left(\frac{\partial P_q}{\partial y}\right)^{-1}\delta P_h \;:\; y = 1\,. \tag{4.1}$$

We can relate this changing to a constant nuclear potential. Like to write

$$\delta\mu_h = -V_0\,. \tag{4.2}$$

V_0 is the potential for each hadron.

So when the hadron (fermions, like protons or neutrons) join, their density increases $\delta\mu_h > 0$ so their pressure rises $\delta P_h > 0$, therefore there is a negative potential $V_0 < 0$. At low energies this potential prevents them from spreading away.

Now we calculate

$$\delta\mu_h = \frac{N_q V_h}{N_h V_q}\left(\frac{\partial\mu_q}{\partial y}\right)\left(\frac{\partial P_q}{\partial y}\right)^{-1}\delta P_h \;:\; y = 1\,.$$

We use the pressure at low energy (3.7)

$$P_{low}/\mu_0^4 = (9\times 1.9\times 240)^{-1}\,7\pi^2 g_q\left[3\times 3.78x^4+\right.$$
$$+3\times 2\times x^2\left(0.82-1.16y+0.41y^2\right)+3\,(0.08-0.23y+$$
$$+0.25y^2-0.12y^3+0.02y^4)+2yx^2\,(-1.16+0.82y)+$$
$$\left.+y\left(-0.23+0.5y-0.36y^2+0.08y^3\right)\right]$$

and we get

$$\frac{\partial P_q}{\partial y} = -\frac{0.076\times 7\times\pi^2\times g_q\mu_0^4}{240\times 3\times 1.9} : x_c = 0.438\,,\ y=1\,. \quad (4.3)$$

Using the relation

$$\mu_0 = \frac{1}{a_0}\left(\frac{8\times 9\pi}{g_q}\right)^{\frac{1}{3}}$$

we have

$$\mu_0 = 285.15\,\text{Mev}\ \text{ for } g_q = 24 \text{ and } 1/a_0 = 135\,\text{Mev}\ \text{ for } \pi^0$$
$$\mu_0 = 295.7\,\text{Mev}\ \text{ for } 1/a_0 = 140\,\text{Mev}\ \text{ for } \pi^- \text{ and } \pi^+\,.$$

So the chemical potential μ_0 is in the range from 285.15 Mev to 295.7 Mev therefore

$$\frac{\partial P_q}{\partial y} = -6.06\times 10^8\,\text{Mev}^4\ \text{ for } \mu_0 = 285.15\,\text{Mev}$$
$$\frac{\partial P_q}{\partial y} = -7.01\times 10^8\,\text{Mev}^4\ \text{ for } \mu_0 = 295.7\,\text{Mev}\,.$$

Now we try to calculate $\partial\mu_q/\partial y$, according to low energy

$$E_{low} = (1.9\times 240)^{-1}\,7\pi^2\,\mu_0^4\,g_q\,V\,x^4\left[3.78+\right.$$
$$+2x^{-2}\left(0.82-1.16y+0.41y^2\right)$$
$$\left.+x^{-4}\left(0.08-0.23y+0.25y^2-0.12y^3+0.02y^4\right)\right]$$

we can equivalence

$$\mu^4 = \mu_0^4\left(1-\frac{0.23}{0.08}y+\frac{0.25}{0.08}y^2-\frac{0.12}{0.08}y^3+\frac{0.02}{0.08}y^4\right). \quad (4.4)$$

But $\partial\mu/\partial y\to\infty$ when $y\to 1$ so we replace

$$\frac{\partial\mu_q}{\partial y}\to\frac{\mu_{y=1}-\mu_{y=0}}{1-0} = \frac{0-\mu_{y=0}}{1-0} = -\mu_{y=0} = -\mu_0\,.$$

Therefore we have

$$\delta\mu_h = \frac{N_q V_h}{N_h V_q}\left(\frac{\partial\mu_q}{\partial y}\right)\left(\frac{\partial P_q}{\partial y}\right)^{-1}\delta P_h$$
$$= \frac{N_q V_h}{N_h V_q}\mu_0\left(0.09\mu_0^4\right)^{-1}\delta P_h$$
$$= \frac{N_q V_h}{N_h V_q}\left(0.09\mu_0^3\right)^{-1}\delta P_h\,.$$

So we have

$$\delta\mu_h = \frac{N_q V_h}{N_h V_q}\left(0.09\,\mu_0^3\right)^{-1}\delta P_h$$
$$= 4.7\times 10^{-7}\frac{N_q V_h}{N_h V_q}\,\delta P_h\ \text{ for } \mu_0 = 285.15\,\text{Mev}$$

and

$$\delta\mu_h = 4.2\times 10^{-7}\frac{N_q V_h}{N_h V_q}\,\delta P_h\ \text{ for } \mu_0 = 295.7\,\text{Mev}\,. \quad (4.5)$$

We use them to cancel the dark matter and dark energy.

5 The Big Bang

We assume there were two universal phases, high energies massless particles phase (let them be the quarks plasma) and then the massive low energies particles (let them be the hadrons).

The first phase associated with high energy density (drops from infinity to finite), the time of that stage is $\tau: 0\to a_0 = 1/(135-140)$ Mev then the massive hadrons phase begins (the time $t: 0\to\infty$).

In both stages the highest universal expansion must not exceed the light speed, for the first phase, high energies massless quarks phase, the density of the energy is the same in all space points so the universal expansion is the same in every point in the space, we let the speed of that expansion equal the light speed, therefore the Hubble parameter $H(t<a_0)$ of this stage $t<a_0$ is given by (5.2).

To find the Hubble parameter for the massive hadrons phase $H(t>a_0)$, we suggest the geometry transformation (5.3) in which the time $\tau: 0\to a_0$ for the quarks corresponds to the time $t: 0\to\infty$ for the massive hadrons phase. We can relate that change in the geometry to the high differences in the energy densities of the two phases. The phase $\tau: 0\to a_0$ high quarks energy, uniform high energy density, massless, ... The phase $t: 0\to\infty$ the massive hadrons, low energy density, separated particles, ...

Now we try to explain how the universe exploded and expanded, we start from our assumptions we made before and find the Hubble parameter and try to find the dark energy and matter. We found that the quarks expand to the length $a_0 = 1/(135-140)$ Mev then the hadrons appear instead.

We assume that the universe was created in every point in two dimensional space XY then the explosion in the Z direction. That is by the quarks, in each point in the XY plane the quarks were created and then they expanded in each point XY to the length a_0 then the explosion in the Z direction, the result is the universe in the space XYZ. There was no universal explosion in the XY plane, the universal explosion was only in the Z direction, in the plane XY there was extension due to the quarks expanding from $r=0$ to $r=a_0 = 1/(135-140)$ Mev the plane XY was infinity before the quarks expansion and it is infinity after that expansion, what happened is an increase

in the number of the XY points, then the explosion in the Z direction. We assume both expansion (XY and Z) occurred with the light speed c.

To find the lost matter, dark matter and dark energy, we use the relation (4.5) we found before:

$$\delta\mu_h = \frac{N_q V_h}{N_h V_q}\left(0.09\,\mu_0^3\right)^{-1}\delta P_h$$
$$= 4.7\times 10^{-7}\,\frac{N_q V_h}{N_h V_q}\,\delta P_h \ \text{ for } \mu_0 = 285.15\,\text{Mev}$$

and

$$\delta\mu_h = 4.2\times 10^{-7}\,\frac{N_q V_h}{N_h V_q}\,\delta P_h \ \text{ for } \mu_0 = 295.7\,\text{Mev}\,.$$

Here we relate this changing in the pressure δP (independent of time) to the hadrons condensation process to form the nucleuses, where the global pressure $\delta P = \delta P_h$ dropped extremely due to the nuclear attractive potential (make it the nuclear binding energy) $V_0 = (-7-8)$ Mev [3]. This pressure δP_h remains contained in the nucleuses, but globally is not visible.

So there is hidden global pressure δP_h and we have to include that problem in the Friedman equations solutions, we notice that the nuclear attractive potential leads to increasing in the cooled hadrons densities. Therefore the decreasing in the hadrons pressure associated with the increasing of their densities (inside the nucleuses). The result is excess in the local energy density, that effects appear in the equations, that is, the matter density appears to be larger than the right energy density. So there is neither dark matter nor dark energy, it is just global and local densities.

We start from the definition of the scale parameter $R(t)$ for the universe expansion, we write [6]

$$ds^2 = -dt^2 + R^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right). \qquad (5.1)$$

We set $k = 0$ flat Universe. Now we try to find the Hubble parameter

$$H(t) = \frac{1}{R(t)}\,\frac{dR(t)}{dt} = \frac{\dot R(t)}{R(t)}\,.$$

There are two phases $t < a_0$ free quarks phase and $t > a_0$ hadrons phase which is the expansion in the Z direction. That means there are two different spacetime geometry, $t < a_0$ and $t > a_0$.

In the first phase $\tau = t < a_0$ the expansion is the same in all space points, so the expansion velocity

$$\frac{dR_1}{dt} = \dot R(t)\,r$$

is the same in all space points and equals the light speed $c = \hbar = 1$ here, so

$$1 = \dot R(t) r : \ t < a = a_0\,.$$

Therefore

$$\dot R(t) = \frac{1}{r} : \ t < a = a_0\,.$$

So we can write

$$R(t) = \frac{t}{r} : \ t < a = a_0\,.$$

So the Hubble parameter becomes

$$H(t) = \frac{\dot R(t)}{R(t)} = \frac{1/r}{t/r} = \frac{1}{t} : \ t < a = a_0\,. \qquad (5.2)$$

Now we want to find the Hubble parameter in the phase $t > a_0$ low energy phase. Actually when the quarks expand from $r = 0$ to $r = a \to a_0$ there will be infinity points expanding, so infinity expanding distance in XY space, but the expansion cannot exceed the light speed $c = 1$ therefore an explosion occurs in the Z direction, so the universal explosion. Therefore the time $t = \tau : 0 \to a_0$ for the free quarks phase will associate with $t : 0 \to \infty$ for the universal expansion, so we make the geometry transformation

$$t = \frac{-c_0}{\tau - a_0} : \ \tau < a_0\,. \qquad (5.3)$$

c_0 is constant, we can relate that relation to a difference in spacetime geometry. That means if the quarks space $r < a_0 = 1/(135-140)$ Mev is flat, so the hadrons space is not, it is curved space, where we live. It is convenient to consider the quarks space ($r < a_0$ large energy density) is curved not our space (low energy density).

Now we can find the Hubble parameter for the universe $t : 0 \to \infty$. We can find the Hubble parameter $H(t > a_0)$ for the geometry $t : 0 \to \infty$ from $H(t < a_0)$:

$$H(\tau < a) = \frac{1}{R(\tau < a)}\,\frac{dR(\tau < a)}{d\tau}$$
$$= \frac{1}{f(r,\theta,\varphi)R(t > a)}\,\frac{d}{d\tau} f(r,\theta,\varphi)R(t > a)\,.$$

We set the geometry transformation

$$R(\tau < a) = f(r,\theta,\varphi)\,R(t > a)$$

so

$$\frac{1}{\tau} = \frac{1}{f(r,\theta,\varphi)R(t > a)}\,\frac{d}{d\tau} f(r,\theta,\varphi)R(t > a)$$
$$= \frac{1}{R(t > a)}\,\frac{d}{d\tau}R(t > a) \qquad (5.4)$$

or

$$\frac{1}{\tau} = \frac{1}{R(t > a)}\,\frac{dt}{d\tau}\,\frac{d}{dt}R(t > a)$$
$$= \frac{dt}{d\tau}\,\frac{1}{R(t > a)}\,\frac{d}{dt}R(t > a) = \frac{dt}{d\tau}H(t > a)\,.$$

Using the geometry transformation

$$t = \frac{-c_0}{\tau - a_0} : \tau < a \to a_0 ,$$

we have the Hubble parameter of the low energy density of the cold Universe

$$H(t > a_0) = \frac{1}{R(t > a_0)} \frac{d}{dt} R(t > a_0)$$
$$= \frac{c_0}{t(a_0 t - c_0)} = \frac{1}{t\left(\frac{a_0}{c_0} t - 1\right)} = \frac{1}{t\left(c'_0 t - 1\right)}$$

where c'_0 is constant.

The Friedman equations can be written, for $k = 0$, like [6]

$$3\frac{\dot{R}^2(t)}{R^2(t)} = 8\pi G_N \rho + \Lambda \qquad (1)$$
$$-\frac{\ddot{R}(t)}{R(t)} + \frac{\dot{R}^2(t)}{R^2(t)} = 4\pi G_N(\rho + p) \qquad (2) \qquad (5.5)$$
$$\frac{d}{dt}(\rho + \delta p) = -3(\rho + p)\frac{\dot{R}(t)}{R(t)} . \qquad (3)$$

To control (or cancel) the dark matter and energy, we make the transformations in the Friedman equations which keep the Hubble parameter unchanged

$$3\frac{\dot{R}^2(t)}{R^2(t)} = 8\pi G_N(\rho + \delta P) + \Lambda - 8\pi G_N \delta P \qquad (1')$$
$$-\frac{\ddot{R}(t)}{R(t)} + \frac{\dot{R}^2(t)}{R^2(t)} = 4\pi G_N(\rho + \delta P + p - \delta P) \qquad (2') \qquad (5.6)$$
$$\frac{d}{dt}(\rho + \delta p) = -3(\rho + \delta P + p - \delta P)\frac{\dot{R}(t)}{R(t)} . \qquad (3')$$

So we have (for same Hubble parameter we had before)

$$\rho' = \rho + \delta p_h$$
$$p' = p - \delta p_h$$
$$\Lambda' = \Lambda - 8\pi G_N \delta p_h = 0 .$$

For the universal nuclear condensation, we assume the universal change $\delta\rho = \delta P = \delta p_h > 0$ is independent of the time.

We can say ρ', p' and $\Lambda' = 0, P' = 0$ are for the located matter, when the hadrons are cooled, they condense and locate in small volumes with high matter density, because of the strong nuclear attractive interaction, so their pressure decreases extremely $P' \approx 0$. That pressure is contained (hidden) in the nucleus. It is like to condense a gas with certain mass m and fixed volume V, the density m/V is the same before and after the condensation, but the real density of the produced liquid is not. Like that we consider ρ the right matter ρ_{matter} and the problems; the increasing $\rho' = \rho + \delta p_h$ and $\Lambda \neq 0$ are because of the phase changing.

We set $\rho' = \rho(t)$ and solve the two equations:

$$-\frac{\ddot{R}(t)}{R(t)} + \frac{\dot{R}^2(t)}{R^2(t)} = 4\pi G_N \rho(t) \qquad (2')$$
$$\frac{d}{dt}(\rho + \delta p) = \dot{\rho}(t) = -3\rho(t)\frac{\dot{R}(t)}{R(t)} \qquad (3')$$

using the Hubble parameter

$$H(t) = \frac{1}{R}\frac{dR}{dt} = \frac{1}{t\left(c'_0 t - 1\right)} \qquad : \quad t > a_0 .$$

From (3') we have

$$-\frac{1}{3}\frac{R(t)}{\dot{R}(t)}\dot{\rho}(t) = \rho(t)$$

so (2') becomes

$$-\frac{\ddot{R}(t)}{R(t)} + \frac{\dot{R}^2(t)}{R^2(t)} = -\frac{4\pi G_N}{3}\frac{R(t)}{\dot{R}(t)}\dot{\rho}(t) .$$

This equation becomes

$$\frac{\dot{R}(t)}{R(t)}\left(-\frac{\ddot{R}(t)}{R(t)} + \frac{\dot{R}^2(t)}{R^2(t)}\right) = -\frac{4\pi G_N}{3}\dot{\rho}(t)$$

or

$$H(t)\left(-\frac{\ddot{R}(t)}{R(t)} + H^2(t)\right) = -\frac{4\pi G_N}{3}\dot{\rho}(t) .$$

Using

$$\frac{d}{dt}\frac{\dot{R}(t)}{R(t)} = \frac{\ddot{R}(t)}{R(t)} - \frac{\dot{R}^2(t)}{R^2(t)}$$

we get

$$H(t)\frac{d}{dt}H(t) = \frac{4\pi G_N}{3}\dot{\rho}(t) \to \frac{1}{2}H(t)^2 = \frac{4\pi G_N}{3}(\rho(t) - \rho_0) .$$

For finite results we put $\rho_0 = 0$ so

$$\frac{1}{2}H(t)^2 = \frac{4\pi G_N}{3}\rho(t) .$$

Now we calculate the contributions of the vacuum energy to the total energy using the cosmological constant Λ' from (1')

$$\Omega_{\Lambda'} = \frac{\rho'_\Lambda}{\rho_c} = \frac{\Lambda'}{3H^2} = \frac{3H^2 - 8\pi G_N \rho(t)}{3H^2}$$
$$= 1 - 2\frac{4\pi G_N}{3H^2}\rho(t) = 1 - 2\frac{1}{H^2}\frac{1}{2}H(t)^2 = 0$$

with the critical energy density

$$\rho_c = \frac{3H^2}{8\pi G_N} .$$

So the vacuum energy density is canceled, and the total energy is the matter energy $\Omega_{matter} = 1$ so $\rho(t)/\rho_c = 1$. Here

$\rho(t) = \rho_c$ is $\rho(t) = \rho' = \rho_{matter} + \delta p_h$, so $\rho(t)$ is higher than the right matter ρ_{matter}.

Now we see if this relation is satisfied or not. We use the global change on the pressure $\delta p = \delta p_h > 0$ which we derived in (4.5):

$$\delta\mu_h = \frac{N_q V_h}{N_h V_q}\left(0.09\mu_0^3\right)^{-1}\delta P_h$$
$$= 4.7\times 10^{-7}\,\frac{N_q V_h}{N_h V_q}\,\delta P_h \ \text{ for } \mu_0 = 285.15\,\text{Mev}$$

and

$$\delta\mu_h = 4.2\times\ 10^{-7}\,\frac{N_q V_h}{N_h V_q}\,\delta P_h \ \text{ for } \mu_0 = 295.7\,\text{Mev}\,.$$

Now we try to find V_q/V_h the quarks volume $V_q = S\,d_q$ and the hadrons volume $V_h = S\,d_h$ as shown in Fig. 12 where the

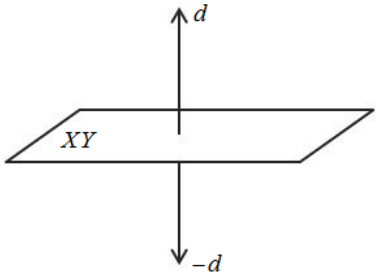

Fig. 12: The universal explosion in Z direction starting from XY flat.

universal explosion is in the $Z = d$ direction. If we assume the explosion speed is the same for both hadrons and quarks, light speed $v = c = 1$, so for the quarks

$$H_q(t) = \frac{\dot R(t)_q}{R(t)_q} = \frac{v_q}{d_q} = \frac{1}{d_q}\,.$$

For the hadrons

$$H_h(t) = \frac{\dot R(t)_h}{R(t)_h} = \frac{v_h}{d_h} = \frac{1}{d_h}$$

therefore

$$\frac{V_q}{V_h} = \frac{S\,d_q}{S\,d_h} = \frac{d_q}{d_h} = \frac{H_h}{H_q}\,,$$

H_h is the universal Hubble parameter, today is

$$H = 71\,\text{km/s/mpc} = 2.3\times 10^{-18}\,s^{-1}$$
$$= 2.3\times 10^{-18}\times 6.58\times 10^{-22}\,\text{Mev} = 151.34\times 10^{-41}\,\text{Mev}\,.$$

The quarks Hubble parameter $H_q = 1/\tau\to 1/a_0 = (135 - 140)$ Mev. So we have (for 135 Mev)

$$\frac{H_h}{H_q} = \frac{151.34\times 10^{-41}\,\text{Mev}}{135\,\text{Mev}} = 1.127\times 10^{-41}\,.$$

Therefore

$$\frac{V_q}{V_h} = \frac{H_h}{H_q} = 1.127\times 10^{-41}\,.$$

We set $\delta\mu_h = -V_0 = (7 - 8)$ Mev the nuclear potential (nucleon binding energy). Therefore, from (4.5), we have

$$\delta\rho = \delta P_h = -\frac{N_h}{N_q}\times 1.127\times 10^{-41}\times\frac{-V_0}{47}\times 10^8\,\text{Mev}^4$$
$$\text{for}\ \frac{1}{a_0} = 135\,\text{Mev} : \mu_0 = 285.15\,\text{Mev}$$

and

$$\delta\rho = \delta P_h = -\frac{N_h}{N_q}\times 1.087\times 10^{-41}\times\frac{-V_0}{42}\times 10^8\,\text{Mev}^4$$
$$\text{for}\ \frac{1}{a_0} = 140\,\text{Mev} : \mu_0 = 295.7\,\text{Mev}\,.$$

For $N_h/N_q = 1/5$, like the interaction $P^+ + \pi^- \to n$ the neutron n appears to have five quarks, that is acceptable according to the fields dual behavior. Therefore

$$\delta\rho = \delta P_h = -\frac{1}{5}\times 1.127\times 10^{-41}\times\frac{-7}{47}\times 10^8\,\text{Mev}^4$$
$$= 335.7\times 10^{-37}\,\text{Mev}^4$$
$$\text{for}\ \mu_0 = 285.15\,\text{Mev and}\ V_0 = -7\,\text{Mev}$$

and

$$\delta\rho = \delta P_h = -\frac{1}{5}\times 1.087\times 10^{-41}\times\frac{-8}{42}\times 10^8\,\text{Mev}^4$$
$$= 414\times 10^{-37}\,\text{Mev}^4$$
$$\text{for}\ \mu_0 = 295.7\,\text{Mev and}\ V_0 = -8\,\text{Mev}.$$

So the change $\delta\rho = \delta P_h$ is in the range:

$$\text{from}\ 335.7\times 10^{-37}\,\text{Mev}^4\ \text{to}\ 414\times 10^{-37}\,\text{Mev}^4\,.$$

Therefore the visible matter is in the range

$$\text{from}\ \rho_{matter} = \rho_c - \delta p_h = 335.7\times 10^{-37}\,\text{Mev}^4$$
$$\text{to}\ \rho_{matter} = \rho_c - \delta p_h = 414\times 10^{-37}\,\text{Mev}^4\,.$$

For the critical energy $\rho_c = 406\times 10^{-37}$ Mev^4 the visible matter is in the range

$$\text{from}\ \rho_{matter} = 0\ \text{to}\ \rho_{matter} = 70\times 10^{-37}\,\text{Mev}^4\,.$$

The right baryonic matter energy density is

$$\rho_b = 4.19\times 10^{-31}\,\text{g/cm}^3 \approx 17.97\times 10^{-37}\,\text{Mev}^4$$

which belongs to the range 0 to 70×10^{-37} Mev^4. We can control this and have

$$\rho_{matter} = \rho_c - \delta P_h = 406\times 10^{-37} - \delta P_h = 17.97\times 10^{-37}\,\text{Mev}^4$$

by finding r:

$$140r + 135(1 - r) = \frac{1}{a_0}$$

where $1/a_0$ satisfies

$$406 \times 10^{-37}\,\text{Mev}^4 - \delta P_h = 17.97 \times 10^{-37}\,\text{Mev}^4.$$

For $1/a_0 = 136.8\,\text{Mev}$ (we used it in (3.17) to have $m_f \approx 938\,\text{Mev}$), the chemical potential becomes $\mu_0 = 288.95\,\text{Mev}$. And with $V_0 = 7.776\,\text{Mev}$ we get

$$\begin{aligned}\delta\rho = \delta P_h &= 335.7 \times 10^{-37} \times \left(\frac{288.95}{285.15}\right)^3 \times \frac{7.776}{7}\,\text{Mev}^4 \\ &= 388 \times 10^{-37}\,\text{Mev}^4\,.\end{aligned}$$

The matter density becomes

$$\begin{aligned}\rho_{matter} &= 406 \times 10^{-37}\,\text{Mev}^4 - 388 \times 10^{-37}\,\text{Mev}^4 \\ &= 17.9 \times 10^{-37}\,\text{Mev}^4\,.\end{aligned}$$

which is the right matter (global visible matter density). Therefore we can control the dark matter and dark energy. We can cancel them.

Note that not all of those ideas are contained in the references.

Submitted on March 16, 2016 / Accepted on March 20, 2016

References

1. Peskin M. E. and Schroeder D. V. An Introduction to Quantum Field Theory. Westview, Boulder, 1995.
2. Srednicki M. Quantum Field Theory. Cambridge University Press, Cambridge, 2006.
3. Martin B. R. Nuclear and Particle Physics. University College, London, 2006.
4. Brown L. S. Quantum Field Theory. Cambridge University Press, Cambridge, 1994.
5. Hooft G. Introduction to String Theory, version 14-05-04. Institute for Theoretical Physics, Utrecht University and Spinoza Institute, Netherlands, 2003–2004.
6. Blau M. Lecture Notes on General Relativity. Institut für Theoretische Physik, Universität Bern, CH-3012 Bern, Switzerland, 2012. http://www.blau.itp.unibe.ch/Lecturenotes.html.

On Quantization and the Resonance Paths

Jacques Consiglio

52 Chemin de Labarthe, 31600 Labastidette, France
E-mail: Jacques.Consiglio@gmail.com

We use a mass-resonance equation to analyze the known elementary particles mass spectrum; we first show that masses and charges are quantized together and all couplings are geometry of movement. Next, the long-expected connection between gravitation and the rest of physics appears as we deduce and compute from the equation parameters the resonance corresponding to the reduced Planck mass. In this way, quantum fields and general relativity can be emergent theories where the natural law is unique.

It is in the admission of ignorance and the admission of uncertainty that there is a hope for the continuous motion of human beings in some direction that doesn't get confined, permanently blocked, as it has so many times before in various periods in the history of man. R.P. Feynman.

1 Introduction

In a celebrated paper, Dirac [8] showed that the existence of magnetic poles and quantum mechanics imply symmetrical quantization of magnetic and electric charges. This is the very first attempt to explain the observation of a universal charge quantum. Since then other theories were produced in which the magnetic charge differs. But even though charges have definite symmetry nothing imposes the charge ratio; namely the fine structure constant α.

It is often believed that the standard model (SM) of particles physics is part of a wider theory in which its free parameters are calculable — but possibly free in essence or accepting multiple solutions. One can see the seeds of this line of thoughts in Dirac's quantization: once the idea is extended to all fields, it may structure the logical constraints in such a manner that the full set of equations can be solved. Such result is expected in super-symmetry and string theory.

However, we must remind that we discuss *the parameters of a theory, not a-priori of nature*. At the other extreme, assume quantum theory incomplete or not fully understood, a possibility exists that all *known* parameters are already calculable from known physics. If so, it may be possible to decode some field characteristics directly from known data. At present time, the only rich group of parameters is the elementary particles mass spectrum as we know 12 samples, and it may be enough to understand its underlying structure.

In short, and in a general manner:

- Assume the 12 known masses correspond to solutions of a set of unknown equations.
- In the most favorable case, if no other mass exists (or close enough) all degrees of freedom are used.
- Hence it may be possible to find or approach the equations and the structure of the solution.

The approach is subtler and a lot more risky than any other since instead of building on theoretical knowledge we assume ignorance — and we do not know what we do ignore.

The object of this paper is to prove the existence of a solution, probably unique, and one of the equations in which the solution is visible. One can infer its validity in two manners; firstly by its agreement with phenomenology, and secondly, by its logical coherence, compactness and simplicity.

In a suite of papers [3, 4], we showed how the mass spectrum is structured. We found firstly that the elementary particles mass obey a simple equation, which is geometrical and based on integral resonances; secondly, two coupling constants (including α) are used in the equation while we find no specific couplings related to the SM weak and Higgs fields as they use only specific geometrical degrees of freedom; thirdly, all calculi and equations are compatible with a simple form of compositeness. On this basis, we showed [5] that the electron and muon magnetic moment anomalies can be computed from the equation parameters with no use of QED.

In the next sections, we first repeat the main demonstrations, fix some errors, and then discuss the results and implications; since the mass equation is geometrical, its use of coupling constants and the manner they combine imply that they are also geometrical; we deduce that they correspond to resonance paths and find or approach the related equations. In this way, the field is geometrically self-quantized and has no free parameter related to energy. The same applies to gravitation since, using Wheeler-Feynman absorber equations, we deduce and compute its coupling (and the reduced Planck mass) from the constants and integral resonances used in the mass equation. In this way this mass-resonance theory is linked to gravitation and cosmology; it needs no dark matter and no big bang but comes with a constant linear expansion and energy creation.

We shall use measurement data and constants from CODATA 2014 or the Particle Data Group 2014 except where mentioned. The point is of importance considering the precision reached with leptons masses, anomalies, and α. The reader should keep in mind that the initial study used older values which imposed no difference to the model.

2 Deriving a mass equation

De Broglie [2] imagined a stationary wave of length $h\,c/E$ which relativistic transformation gives a phase wave of length h/p. This is the origin of the wave equations of quantum mechanics. The question of the nature of those waves is still open; in this section, we imagine how a stationary wave can be born *and ring*; then we predict some characteristics of the resonances that we shall later use as verification.

We assume that the wave is the physical exchange at the origin of mass. Energy exchange is momentum, and it gives a pressure field that "cages" the particle charges and some associated self-energy. The initial idea is similar to the Poincaré stress [11] though not identical as we split the particle.

Roughly speaking, we cage a permanent photon-like current in a box also made of currents and we guess that the box and the charge quantize each other. Assume the box size universal, it is sufficient to use a length 1. In the one dimensional case, the pressure is a simple force, and resonance implies an integral number M such that we have:

$$m = \mu + X\,M,$$

where m is the particle mass and X is a universal constant. The quantity μ represents a massless self-energy that necessarily propagates, and it implies a double resonance. Hence the resonance corresponds to a product $M = N\,P$:

$$m = \mu + X\,N\,P.$$

In the 1-dimensional case, we should have $N = P$ corresponding to identical inbound and outbound currents, but we shall need a more general equation and then we use a product. In a wave representation, it represents the number of times the inbound and outbound wave crests hit each other in a universal period of time or within a definite length.

Caging a massless particle requires symmetry, a force that opposes the particle charge to the pressure field, that is precisely the resonance N P and the self-energy μ. There must be a residual distance $d \neq 0$ between the first resonance wall and the current μ at which the force applies. It gives:

$$m = \mu + \frac{X}{d + \dfrac{1}{N\,P}}.$$

Now the distance d should also depend on N and P because energy comes from the distance (d + *1/NP*) which is equivalent to a potential. A potential is quantized and $1/NP$ is already quantized as it comes from $XNP = XM$. Then we use $d = KD$, with K an integral number and D a length. Last, in three dimensions we get a cube:

$$m = \mu + \frac{X}{\left(K\,D + \dfrac{1}{N\,P}\right)^3}. \tag{2.1}$$

The equation has 6 degrees of freedom that can be reduced to 5 by division by X or μ and give unit-less quantities.

Now let us discuss the equation geometry; contrary to the one-dimensional case, we have more degrees of freedom in the resonance and the paths associated to N and P can be radial or circular; here we can use group theory arguments:

— Case 1: A double radial resonance. It needs identical inbound and outbound waves, then $N = P$, giving a stationary wave. Except for the cube, it is identical to the 1-dimensional case then it should address leptons and U(1), and also the Poincare stress in which case we should have $KD > 0$, with K increasing with mass as $1/NP$ reduces since the leptons charges are identical.

— Case 2: A double circular resonance: The resonance geometry is conserved when we invert rotation axis; hence it must be identified to SU(2) and by symmetry $N = P$. But we must change (2.1) with $X \to X/k\,\pi$ with k a constant integral number; this is because compared to the first case even though the resonance is circular the pressure is still applied to its geometrical center. The equation becomes:

$$m = \mu + \frac{X}{k\,\pi\left(K\,D + \dfrac{1}{N\,P}\right)^3}.$$

It addresses massive bosons, which role in nature is to carry interactions. They are similar to a photon and we must integrate to X the term μ (that would be an intrinsic mass). Therefore we will compute their masses (index b) comparatively to the full electron mass (index e) as follows:

$$\frac{m_b}{m_e} = \frac{\left(\dfrac{1}{N_e P_e} + K_e D_e\right)^3}{k\,\pi\left(\dfrac{1}{N_b P_b} + K_b D_b\right)^3}. \tag{2.2}$$

- Case 3: A mixed resonance. It includes both symmetries U(1) and SU(2), it is then SU(3) and this case addresses quarks. If D is related to the strong force and asymptotic freedom ($\approx$ inverse to the Poincare stress) we should have $KD < 0$, ideally constant. It implies $N \neq P$ with a geometrical constraint between π, N and P since a phase lock between the two paths must exist; it requires to squaring a circle, then logically we should get approximate relations like:

$$N\,P\,\pi \approx \textit{an integral number}, \tag{2.3}$$

If the logic above is valid, it follows that particles distant interactions are a manifestation of the resonance; hence we should find relations between the resonance numbers (N, P) and the known symmetries, and also between some coupling constants and the non-integral values of D, X, and μ. *De facto, and most importantly, we cannot understand mass and charge quantization separately*.

3 Massive elementary particles resonances

In this section, we shall fit the equation parameters to all known elementary particles masses; since the equation is related to symmetry, the natural strategy is to proceed by groups (leptons, quarks, massive bosons). We shall assume X universal and μ specific to leptons (where enough precision exists) and, since D addresses forces, it must be group-dependent.

Recall also that a number of relations must be verified by the fit; they can be used as verification of the geometrical constraints imposed by symmetry and by the equation.

3.1 Leptons

The Table 1 shows charged leptons resonances. It uses very small numbers, we get $N = P$ as expected. The equation parameters are given hereafter:

$$\mu = 241.67661953\,\text{eV},$$
$$D_e = 0.0008532218937, \tag{3.1}$$
$$X = 8.1451213299073\,\text{KeV}.$$

Table 1: Electron, muon, tau in MeV/c^2.

–	P = N	K	Computed	Measured
e	2	2	0.510 998 9461	0.510 998 9461(31)
μ	5	3	105.658 3752	105.658 3745(24)
τ	9	5	1 776.84	1 776.82(16)

Using α, the fine structure constant, we define a new constant that will be used later:

$$A_S = D_e/\alpha \approx 0.11692, \tag{3.2}$$

which name A_S is chosen for its value is reminiscent of the strong force coupling.

The values in (3.1) can be tuned so that all masses match exactly regardless of uncertainty; instead those values have been chosen to compute exactly the electron mass and magnetic moment anomaly (assuming the related equations developed later are good-enough for such precision).

3.2 Quarks

Using X and μ constant from (3.1) the quarks resonances are shown Table 2 (masses in the natural scheme) where a regular pattern is obvious.

As expected, the parameter D is slightly different from (3.1) to compute those masses:

$$D_q = D_e(1 + \alpha) = A_S\,(\alpha + \alpha^2). \tag{3.3}$$

Using D_e like for leptons gives the top mass out of range $\approx$ 167 GeV, and then a difference with leptons exists. Quarks masses are no more published in the natural scheme; the estimates used in Table 2 are dated 2011 except for the top [18], see also [19].

We get $N \neq P$ as expected; P and K are constant which is surprisingly simple. The constancy of $K = -6 < 0$ is reminiscent of asymptotic freedom and then also agrees with a connection between D_e and α_s. Note that varying K by ± 1 gives computed quarks masses out of uncertainty range for the four heaviest.

Table 2: Quarks resonances in MeV/c^2.

–	P	N	K	Computed	Estimate
u	3	2	–6	1.93	1.7 – 3.1
d	3	19/7	–6	5.00	4.1 – 5.7
s	3	7	–6	106.4	80 – 130
c	3	14	–6	1,255	1,180 – 1,340
b	3	19	–6	4,285	4,130 – 4,370
t	3	38	–6	172,380	172,040 ±190 ± 750

The approximate relations with $N\,P\,\pi$ (2.3) are verified for the second and third generations; they are:

$$c, s : 7 \times 3\pi \approx 65.97 \approx 66/1.0004025,$$
$$t, b : 19 \times 3\pi \approx 179.07 \approx 179 \times 1.0003954.$$

We also notice that between 1 and 19 no other integral numbers come close to verifying (2.3).

It is interesting that the multiplication of N by 2 in the second and third generations corresponds to the difference in electric charges $(1/3, 2/3)$ as it links mass and charge quantization. For the first generation the down quark needs a fraction $N = 19/7$ which is barely acceptable, and we notice that the relations with (2.3) match with $2\,\pi$ for the d and also indirectly for the u instead of $3\,\pi$ for the four heavier quarks.

Those particularities may relate to quarks mixing, which we see in the fraction $19/7 = 38/14$, and the same logic for u also holds since $2 = 38/19 = 14/7$.

$$u : 2 \times 3\pi \approx 19/1.008,$$
$$d : (19/7) \times 2\pi \approx 17 \times 1.0032.$$

Hence something unique happens to the u and d.

3.3 Massive Bosons

We assume that the $W^{\pm}$, Z^0 and H^0 acquire their masses from the same geometry; recall that we only have three geometries (or mechanisms) and then we cannot address the weak force bosons and the H^0 separately. Using (2.2), it corresponds to the same resonance, that is on the circular path we must have $N = P = constant$, and only the radial K varies (though this is not exact since we shall later find a slight difference).

A factor $k\,\pi$ at the denominator of (2.2) is needed since the resonance is supposed circular, but we do not find a perfect fit with k integral. We need a factor $k \approx 1$; it seems at first that we add a degree of freedom but we shall show that it is a geometrical constraint.

The analysis of those masses is iterative and leads to important reasoning which is repeated hereafter in details. In practice:

- The empirical fit gives the resonances, which are $N = P = 12$, and $K = -2, -7, -19$ for the $W^{\pm}$, Z^0 and H^0 respectively. The weak force bosons come in range but the error on the H^0 is 1 GeV. Those numbers immediately suggests the same underlying geometry as quarks and maybe leptons, then the same field combining potentials expressed by D_e and α.
- The empirical value of D for massive bosons is first approximated as $D_b \approx \alpha^2(1+A_S/2-A_S^2/6)$; it suggests an interaction term that depends on α and D_e; the former is known and the later estimated with precision.
- The expression $[D_e(1+\alpha)]^2 = \alpha^2(1+2\,A_S+A_S^2)$ is similar and may give $D_b \approx \alpha^2(1+A_S/2-A_S^2/6)$ depending on the effective algebra. (Doubling the forces divides the distance, then $2A_S \rightarrow A_S/2$, and the term $-A_S^2/6$ fits with the $K = -6$ in table 2.) .

On this basis we may have enough information to model the interaction; the equations (3.2 – 3.3) suggest:

- Two types of charges corresponding to the mass $\mu : E$ and C ($\approx$ electric and color) on which D depends.
- A free field (charges X), and the pressure is given by interactions: $X\times X$, $E\times X$, and $C\times X$, hence D_b includes 3 terms, but its expression is incomplete as we do not yet compute all masses with precision.

Now we shall complete the reasoning, compute the predicted bosons masses, and compare to experimental data.

Classification and immediate identification gives Table 3. It shows that each individual interaction adds a piece of coefficient in D_b — like simple potentials adding or subtracting. But we can only compute a radial distance (which gives a radial strength), not the orientation of the force which can be symmetry-dependent as we discuss rotations.

Table 3: Classification and minimal interpretation of the coefficients.

–	D	Coeff	Interaction	Interpretation/logic
1	D_e	αA_S	$X\times E$	*Leptons*
2	D_q	αA_S	$X\times E$	*Leptons* $\rightarrow$ *Quarks*
3	D_q	$\alpha(\alpha A_S)$	$X\times C$	*Quarks Charge*
4	D_b	α^2	$X\times X$	–
5	D_b	$\alpha(\alpha A_S)/2$	$X\times C$	*Quarks* $\rightarrow$ *Bosons*
6	D_b	$(\alpha A_S)^2/6$	$(X\times E)^2$	*Leptons* $\rightarrow$ *Bosons*

The important point in this table is that quarks charges resume to $X\times C = X\times(X\times E)$, and the coefficient 1/2 line 5 implies two distinct charges (augmenting the force and then reducing the distance). Interpretation details are given hereafter (referring to the line of the Table 3) and lead to understanding.

Leptons — Line 1; charge E.

- $X\times E \rightarrow \alpha A_S$: There is only one elementary interaction; it just gives us its coefficient.

Quarks — Lines 2 and 3; charges E and C.

- $X\times E \rightarrow \alpha A_S$: Same as electrons, and independent of the quark electric charge.
- $X\times C = X\times(X\times E) \rightarrow \alpha(\alpha A_S)$: This is a different interaction; it is not a new kind of charge but it has the same nature and quantum as X.

Massive Bosons — Lines 4, 5, and 6: charges E and C.

We found the same coefficients for the $W^{\pm}$ and the Z^0. One is electrically neutral but not the other. Still, we find coefficients related to electricity and color charge, and then those bosons are made of two fractional electric charges and their two color charges (as we shall see the term charge is abusive here). Then it is:

- $X\times X \rightarrow \alpha^2$: The interaction of two charges X gives a distance α^2. This is the main force on the circular path that other interactions will impact — they are secondary forces or loops impacting this path.
- $X\times C = X\times(X\times E) \rightarrow \alpha(\alpha A_S)/2$: The coefficient $\alpha(\alpha A_S)$ comes with quarks color charge; it also shows that the charges of a weak force boson are equivalent to that of two quarks, and different of that of a lepton. Increasing the force by a factor 2 reduces the length proportionally; thus the factor 1/2.
- $(X\times E)\times(X\times E) \rightarrow -(\alpha A_S)^2/6$: This coefficient corresponds to the effect of the main resonance on separate electric charges. We recognize $D_e = \alpha A_S$ from leptons, but 1/6 is new; it is only associated to D_e^2 and this interaction is not present in Tables 1 and 2.

At this point, we understand how the interaction works and we can logically deduce all missing terms in the expression of D_b using α and A_S. For this, we need to complete the series of interaction loops with the field X:

$X\times X\times X \rightarrow -\alpha^4$: Since $X\times X \rightarrow \alpha^2$ positive, and $K < 0$, the force in $X\times X$ is compressive and then this coefficient is scalar (and positive), it increases the compression and then reduces the length: the coefficient is then negative $-\alpha^4$. The next coefficient is positive as it reduces $-\alpha^4$. Similarly, we must add loops indefinitely ($X\times X\times X\times X$ etc.); it gives a simple series converging to $\alpha^2/(1+\alpha^2)$.

Last, each interaction must be augmented with any number of X where the corresponding length is modified depending on its sign; then the coefficient $-A_S^2/6$ is multiplied by $1/(1+\alpha^2)$ and the coefficient $A_S/2$ by $1/(1-\alpha^2)$. The series make a small difference in D_b which is far from negligible when it comes to computing masses. The coefficient D_b for the $W^{\pm}$ and Z^0 is then:

$$D_{WZ} = \alpha^2 \left(\frac{1}{1+\alpha^2} + \frac{A_S}{2(1-\alpha^2)} - \frac{A_S^2}{6(1+\alpha^2)} \right),$$

$$D_{WZ} = 5.62404904 \times 10^{-5}. \tag{3.4}$$

It also reads:

$$D_{WZ} = \frac{\alpha^2}{1+\alpha^2} + \frac{D_e}{2(1-\alpha^2)} - \frac{D_e^2}{6(1+\alpha^2)}.$$

But it cannot be identical for the H^0, firstly because its spin is not 1. Assuming it holds four charges organized in a tetrahedral manner, a tetrahedron has 6 lines of forces, and the last interaction term is six times stronger:

$$D_H = \alpha^2 \left(\frac{1}{1+\alpha^2} + \frac{A_S}{2(1-\alpha^2)} - \frac{A_S^2}{1+\alpha^2} \right),$$

$$D_H = 5.56338664 \times 10^{-5}. \tag{3.5}$$

Or, alternately,

$$D_H = \frac{\alpha^2}{1+\alpha^2} + \frac{D_e}{2(1-\alpha^2)} - \frac{D_e^2}{1+\alpha^2}.$$

It may also include additional loops thru the tetrahedron. The strength of a line linking two charges is $1/6$, it gives the first term A_S^2 in (3.5), but for the H^0 it propagates thru 6 lines of a tetrahedron and it gives $6\,A_S^4$. But it is not a free field, and then it may not need an infinite number of loops. We shall use a one-loop approximation since additional loops makes a small difference ($\approx -10\,MeV$):

$$D_H = \alpha^2 \left(\frac{1}{1+\alpha^2} + \frac{A_S}{2(1-\alpha^2)} - \frac{A_S^2\,(1+6A_S^2)}{1+\alpha^2} \right),$$

$$D_H = 5.55741566 \times 10^{-5}. \tag{3.6}$$

This expression is the only reason here for A_S to be physical since all others uses of this coefficient reduce to D_e.

Now let us come back to the coefficient k in (2.2). In Table 4, we have $N = P$, and then those two resonances have the same orientation with opposite paths, but we find K in $(-2, -7, -19)$ the same numbers as for the quarks N which resonance is mixed.

Consequently, there is, like for quarks (2.3), a geometrical constraint which here is between the length D_b and the circular path π/NP. Taking only the circular path into account and keeping the constraint coming from the radius, D_b should be a divisor of $\pi/NP = \pi/144$, a division that must hold with any K in $-2, -7, -19$. Since all Ks are primes numbers the constraint applies to their product. In this simplified picture (that cannot hold yet) we should have:

$$(\pi/144)/D_b = 2 \times 7 \times 19 \rightarrow \pi/144 = 266\,D_b$$

Now D_b is radial and a 3-sphere volume depends on the cube of its radius. Then we must use $D_b\pi^{1/3}$ on the right hand side; it gives a modified equation that is close to hold:

$$\pi/144 = 266\,D_b\pi^{1/3}.$$

This equation is equivalent to squaring the circle, then we miss the coefficient k which is now a logical geometrical constraint related to phase lock. In (2.2), π is multiplied by k and this equation addresses a volume; hence we must use its cube on the left hand side, and reduce π accordingly on the right-hand side; in this way we get comparable quantities and it gives the geometrical resonance constraint:

$$k^3\,\pi/144 = 266\,D_b\,(\pi/k)^{1/3}. \tag{3.7}$$

Here the interaction term D_b constrains k thru geometry. The two sides of (3.7) represent lengths, and then taking their cube we get volumes verifying:

$$(266\,D_b)^3 = k^{10}\,\pi^2\,(1/144)^3. \tag{3.8}$$

It equates the volume of a 3-cube of edge $266\,D_b$ on the left hand-side to that of a 4-ball ($V^4 = \pi^2\,R^4/2$) divided by half its radius on the right-hand side, where a correction k is needed for cubing the sphere. Here D_b is an interaction term in 4D, k a geometrical wave coherence constraint, and (3.8) links a radial and a circular path in 4D. Now compute from (3.8):

$$(3.4) \rightarrow k_{WZ} = 1.00128565, \tag{3.9.1}$$

$$(3.5) \rightarrow k_H = 0.998033312, \tag{3.9.2}$$

$$(3.6) \rightarrow k_H = 0.997711845. \tag{3.9.3}$$

Using the coefficients above and (2.2), gives the masses in Table 4, where precision is impressive.

Table 4: Bosons resonances in MeV/c^2, H^0 mass in [17].

–	P = N	K	Computed	Measured
$W^{\pm}$	12	−2	80,384.9	80,385 ± 15
Z^0	12	−7	91,187.56	91,187.6 ± 2.1
H^0	12	−19	125,206	125.090 ± 240
H^0	12	−19	125,094	125.090 ± 240

After modeling the interaction we compute the weak force bosons masses in perfect agreement with measurement and it and confirms the validity of our reasoning. We get an effective

unified theory of resonances where the forces compositeness decays from leptons and quarks and this is truly unexpected.

In this table, the last two lines correspond to the equations (3.5 – 3.9.2) and (3.6 – 3.9.3) respectively for D_H and k_H.

We can now better analyze the resonance in Table 4. Consider the length $2 \times 7 \times 19 = 266$. A phase lock between the radial and circular paths and the $K = -7$ and -19 imply two circular path lengths which are $L1 = 2\pi(1 - 7/266)$, and $L2 = 2\pi(1 - 19/266)$. Those are compatible if and only if $A\,L1 = B\,L2$, with A and B integral numbers. We must solve the following equation which solution is trivial:

$$\frac{A \times 2\pi(266-7)}{266} = \frac{B \times 2\pi(266-19)}{266},$$

$$A = 266-19 = 247, B = 266-7 = 259, B - A = 12. \quad (3.10)$$

The resonance number, 12, appears on the left hand side of (3.10); it comes from phase coherence between the circular path and the spots on the radius and we naturally get $N = P = A - B = 12$ which then depends only on K (we use only -7 and -19, but $K = -2$ is not a problem since 12 is even).

Finally all numbers and parameters used in Table 4 appear constrained; the specific degree of freedom used here is just geometry. We have two forces coefficients (α and D_e) and no specific coupling in this sector which is then emergent; this result disagrees with the SM concept and requires unification from below (as opposed to distinct fields).

3.4 Bosons widths

The expression (2.2) is a resonance equation and the computed masses correspond to the poles of the resonances. Then it should be possible to compute widths and then lifetimes; at best, the widths are the size of some working resonance "spots"; it would show that this theory gives the SM weak field. For this we have to understand the phase coherence between multiple paths. Recall that the bosons charges are found interacting and organized in a minimal manner; in 3D, it is a tetrahedron for the H^0 and a simple straight line for the Z^0 and $W^\pm$. For the weak force bosons:

With two circular phases the symmetry is loose, it has some freedom, and on the circular path it suffices that N and P hold on $1/2$ phase to stabilize the resonance. It authorizes a circular phase shift $\pm\pi/12$ which extends or reduces the sphere; with two charges, it gives on the radial part $\Delta K = (\pm 1/2)(1/12) = \pm 1/24$.

In the radial direction, we have 266 slots, and the same reasoning applies; it adds $\Delta K = \pm 1$.

For the H^0, with 4 charges, the symmetry is fully constrained in 3D; N and P hold together: $\Delta K = 1/144$. A tetrahedron has 6 lines of force that can break; hence the width is reduced accordingly $\Delta K = 1/144/6$. Other loops add nothing since a tetrahedron is fully constrained in 3D.

On this basis, the resonance width is the difference in mass ΔM given by (2.2) with respect to the pole in Table 4 when we use $K + \Delta K$ in (2.2) to compute the particle mass $M + \Delta M$. We get:

$W^\pm \to \Delta K = (1 + 1/24) \to \Gamma_W = 2.0857$ GeV, a perfect match with experiment (2.085 ± 0.042 GeV).

$Z^0 \to \Delta K = (1 + 1/24) \to \Gamma_Z = 2.468$ GeV, 1% less than expected (2.4952 ± 0.0023 GeV).

$H^0 \to \Delta K = 1/(144 \times 6) \to \Gamma_H = 4.10$ MeV, which agrees with the SM prediction at 125.09 GeV.

Hence, the widths come straightforwardly from geometry. But the Z^0 width is out of range and this can only be due to the difference in charges with the $W^\pm$ that we have ignored. Reasoning simply:

$W^\pm$: The charges e/3 and 2e/3 (or opposite) repel each other with a force coefficient $2e^2/9$.

Z^0: The charges e/3 and -e/3 (or 2e/3 and -2e/3) attract each other, the force coefficient is $e^2/9$ or $4e^2/9$.

The difference in inner charges between the Z^0 and the $W\pm$ gives a difference in forces which is:

$$\frac{2e^2}{9} + \frac{e^2}{9} = \frac{e^2}{3} \text{ Or : } \frac{2e^2}{9} + \frac{4e^2}{9} = \frac{2e^2}{3}.$$

It implies that the forces cannot be balanced in the same manner for the two bosons. Assuming the $W^\pm$ width computed value is exact, we need an additional term to compute the Z^0 width. Since the forces in the calculus of D_b depend on charges, from the equations above the missing coefficient is $1.5/137$ or $1.5\,\alpha$. It gives:

$Z^0 \to \Delta K = (1 + 1/24 + 1.5/137) \to \Gamma_Z = 2.4946$ GeV,

which agrees with the SM prediction and experimental data. However the experimental precision for the $W^\pm$ and Z^0 widths differ by one order of magnitude; hence this reasoning, which is differential, is risky and non conclusive.

3.5 Resonance terms, analysis and reduction

The resonance terms found in the previous tables (all N and P) reduce to 2, 3, 7, and 19 in the following manner:

Leptons: 2, 7 – 2, and 7 + 2.

Quarks: 3, 7, 2×7, 19, and 2×19, if we omit the u and d where we know from the CKM matrix that mixing is large as compared to the other angles.

Massive bosons: 12 = 19 – 7.

It is remarkable that $7 = 2^3 - 1^3$, and $19 = 3^3 - 2^3$; here it reduces to the three "symmetry numbers" of U(1), SU(2) and SU(3), and their cubes differences. Moreover for all quarks we get $P = 3$, including the u and d, where the polarity appears, meanwhile for leptons it seems that we have the polarity 2 in a mixed manner. In this way the radial paths are based on 2 and 3, while 7 and 19 only come with circular paths (and unstable or mixing particles).

Moreover, the difference in resonance between the electron and the muon and tau relate to the $K = -7$ of the Z^0, while the heavy quarks decays include a factor 2 in charge and resonance which fit the $K = -2$ of the $W^\pm$.

Therefore we get the strong impression that the equation relates to an intricate resonance scheme based only on the SM symmetries — or something close. The simplicity of the reasoning and numerical results suggest that the mass spectrum may be unavoidable, and since it relies on charges it also suggests the absence of free parameters in nature.

3.6 Charges ratios

The results in this section suggest a single field "below" and it is interesting to estimate charges ratios, but we can only compute their radial effect, not the forces orientation; from the analysis of Table 3 the distances building the Ds are in reverse proportions of charges. Then for the electron, $K = 2$, and for quarks, $K = -6$, and $D_e \approx D_q$.

From Table 3 and the different parameters D, taking into account the differences in K, since we have $X \times E \to 2 \times D_e$ for the electron and for quarks $X \times C \to -6\,\alpha\,D_e$; we estimate:

$$\frac{C}{E} = \frac{2\,D_e}{6\,\alpha\,D_e} \to C = \frac{E}{3\,\alpha},$$

which has a clear scent of monopole; importantly, it does not depend on the quark electric charge since the coefficient 3 (from $K = -6$) is constant in Table 2.

In Table 3, we also have $X \times X \to \alpha^2$ and $X \times C \to 6\,\alpha\,D_e$, and then we estimate:

$$\frac{C}{X} = \frac{\alpha^2}{6\,\alpha\,D_e} = 1.4254503 \approx \sqrt{2},$$

which is in the range of 1 and then the same type of charges ($\approx \sqrt{2}$ suggests geometry of the force orientation).

4 Coupling constants

4.1 Introduction

We found two real constants in the expression of the parameters D which represents a length in the equation. In the expressions (3.4 – 3.5 – 3.6) used for D_b those two constants stand on equal grounds. Hence since α is the coupling constant of QED, then D_e is also a coupling constant. It then relates directly to the strong and weak forces couplings (recall that we also have $A_S = D_e/\alpha$ in the range of $\alpha_{S(MZ)}$) and since K appears constant for quarks, D_q should be related to asymptotic freedom. Therefore it seems that the equation addresses a field below with two and only two couplings (neglecting gravitation for now).

Since all resonances are integral (N, P, K) and reduce to a few numbers, it is minimal and elegant to generalize the concept and assume that the field is entirely self-quantizing (or self-constraining) and that quantization is entirely based on geometry and integral numbers; in this way, those two coupling correspond to some counter-resonances ($1/N \to N$ or $N \to N$) and then to constant path lengths (or relative path lengths).

In practice the only known constant integral path length is that of photons for which $r^2 - c^2t^2 = 0$. At the opposite, in special relativity, massive particles obey $r^2 - c^2t^2 = const \neq 0$ which we write $r^2 - c^2t^2 - const = 0$. But now the paths of the resonance define the massive particles — *we mean entirely*; it is a repeat pattern that fits into this equation and it first implies that the path includes a rotation which is around the time axis. Then we guess that D_e (as a length) must be computed from a pseudo-norm like expression of the form:

$$n^2 + m\,\pi^2 - p^2 = D_e^{-2}$$

where the central term introduces a rotation and n, m, and p are expressions based on the resonance terms. Now of course, α must obey a similar pattern and, since α and D_e have distinct but complimentary roles, the expressions giving D_e and α should use resonance terms in a complimentary manner. Last, the bosons resonances are based on 4-dimensional paths; then n, m, and p must be seen as the coordinates of a 4-path which projection on 3-dimensional space gives real numbers.

Because of 4D resonances, we shall suppose that there is no punctual particle or 1D string and that the field is entirely fluid. It implies that some currents propagating in a direction orthogonal to the observable 3-space (possibly back and/or forth in time) are preserving and propagating the characteristics of the particle and we shall abusively denote those "time-currents". In this way the electric field of the electron is seen similar to the effect of a magnetic current propagating forward and/or backward in time with respect to the present. Here the present is seen as the surface of an expanding 4-sphere, but 4D space is assumed preexisting and permanent.

It results in an interesting minimal model where all known massive particles are composites of time-currents:

Leptons:

- e^- : $[\uparrow^-\downarrow^+]$,
- μ^- : $[\uparrow^-\downarrow^+\downarrow^-\downarrow^+]$,
- τ^- : $[\uparrow^-\downarrow^+\uparrow^-\uparrow^+]$.

Quarks:

- t^+ : $[\uparrow^+\downarrow^-\downarrow^+\uparrow^-\uparrow^+]$,
- b^- : $[\downarrow^+\uparrow^-\uparrow^+]$,
- $c+$: $[\uparrow^+\downarrow^-\downarrow^+]$,
- s^- : $[\downarrow^+]$.

Bosons:

- Z^0: $[\downarrow^+\downarrow^-]$,
- $W^\pm$: $[\uparrow^+\downarrow^-]$ and $[\uparrow^-\downarrow^+]$,
- H^0 : $[\uparrow^+\uparrow^-\downarrow^+\downarrow^-]$.

where the notations are trivial for up-time and down-time currents sign and directions (the sign is the current, not the electric charge which, by convention, is inverted for down currents); the apparent electric charge is $2/3$ for an up-time cur-

rent and 1/3 for a down-current (still by convention). Several aspects of the model are of interest:

- The model is based on 4 dimensions of space; it is then coherent with the calculus of the coefficient k used for bosons, but it is also reminiscent of QCD where quarks live in 4 dimensions.
- The difference between the H^0 and the weak bosons is consistent with the calculus of D_{WZ} and D_H.
- All quarks decays consist in a separation of currents where the sum of the produced $W^{\pm}$ boson's current and quark is equal to the currents of the original quark (and of course the picture is reversible).
- The same is valid for leptons decay, but with a Z^0.
- There is no room to make a d quark except by mixing (and the d comes with resonances ratios).
- The notion of time-currents removes the need for particles "inhabiting space". In this way, the concept is minimalist and particularly elegant since, eventually, it must result in self-quantizing movement where we do not need to distinguish space and matter.
- All particles include a down-type current (taking this as strict rule implies mixing for the u and d, and the absence of FCNC). The model agree with Cramer's interpretation of quantum mechanics — though in an almost classical 4-dimensional manner. All particles are connected to and can send information to their past or receive some from their future because a communication channel exists which is the particle itself.

4.2 Coincidences

In this sub-section we discuss three numerical coincidences involving the numerical values found in section 3. In this way, we seek coherence with known but older theory.

4.2.1 Lamb shift, Bethe's equation

Bethe [1] computes the hydrogen Lamb shift; he gets:

$$\Delta E = \frac{\alpha^5 m_e c^2}{6\pi} \ln\left(\frac{m_e^2 c^2}{8.9\,\alpha^2 m_e^2 c^2}\right), \tag{4.1}$$

where m_e is the electron mass; the expression in the logarithm depends on the cutoff and gives a ratio between the electron absorption and self-interaction and then in our model μ and $(m_e - \mu)$ respectively (though according to the mass-equation, self-interaction and absorption may be reversed with respect to QED,) we find:

$$\frac{(m_e - \mu)}{\mu} = \frac{1}{8.8857\,\alpha^2}. \tag{4.2}$$

The relative difference with respect to Bethe's result is 1.6×10^{-3} (or 2×10^{-4} for ΔE) and then μ seems relevant with respect to Bethe's analysis. We notice a similar coincidence:

$$\frac{(m_e - \mu)}{\mu} \approx \frac{\sqrt{2}}{4\pi\,\alpha^2}. \tag{4.3}$$

The relative error in (4.3) is $\approx 1.25\times10^{-5}$. Consequently, since Bethe's paper is seen as the very first step to QED, X and μ should be fundamental quantities directly linked to QED.

4.2.2 The electron mass and spin, rough analysis of the coincidences

A physical action is a product of charges or currents; then we analyze action and not energy. Accordingly, the electron mass comes as a repeated action ($E = h\nu$).

Action is a product that we first write in complex form:

$$\left(G + \frac{i\,e}{2}\right)\left(G - \frac{i\,e}{2}\right) = G^2 + \frac{e^2}{4} \to m_e, \tag{4.4}$$

where e/2 represents the currents, not the apparent charges, and G the resonant component. Now we write (4.4) in quaternion form:

$$\left(G + \frac{i\,e}{2}\right)\left(G + \frac{k\,e}{2}\right) = G^2 - \frac{j\,e^2}{4} + (k+i)\frac{e\,G}{2}. \tag{4.5}$$

Those equations may approach the natural algebra, but the result seems wrong. Still, assume the algebra is broken, (4.4) gives the mass and (4.5) angular momentum:

$$G^2 + \frac{e^2}{4} \to m_e;\ (k\ +\ i)\frac{e\,G}{2} \to \textit{angular momenta}. \tag{4.6}$$

The angular momentum splits into two components on orthogonal axis — which agrees with the idea of time-currents. Then one is the magnetic moment and the other is along the time axis; we will denote the latter "spin". Now we identify the squared charges in (4.6) with the masses in (4.3); it gives:

$$4\pi\,\alpha^2 G^2 \approx e^2\frac{\sqrt{2}}{4}.$$

Substituting G with a Dirac charge, we get $1 \approx \sqrt{(2)}/4\pi$; now multiply each side of this ridiculous result by the Planck constant we get the following correspondence:

$$h \leftrightarrow \sqrt{2}\,\frac{\hbar}{2} = \left|(k+i)\frac{e\,G}{2}\right|, \tag{4.7}$$

which interpretation is obvious: a repeated action h is energy ($E = h\nu$) and it makes the leptons spin and magnetic moment.

4.2.3 The Dirac condition and the parameters X and μ

Dirac [8] analyzes the possibility of existence of magnetic monopoles using quantum mechanics. Based on the mathematical properties of the electron wave function interpreted

as a density of probability of presence, he shows that a magnetic monopole is compatible with the existence of quantum mechanics in Hamiltonian form if and only if the so called Dirac condition is respected:

$$e\,g = \frac{(n\,\hbar\,c)}{2} \rightarrow g = \frac{n\,e}{2\,\alpha}. \tag{4.8}$$

It results in the elegant idea that the existence of magnetic poles fixes the electric charge and conversely.

Now let us assume that the electron wave is a magnetic current; since Dirac's demonstration is based on the "fields of force" acting on the electron wave then magnetic currents acting on electric charges must obey the same condition. But in our model e is an apparent charge (say e_e) and also a sum of time-currents (say e_m) and its monopole (denoted g_m).

Both must be taken into account in the condition as part of the total current; then the condition is:

$$e_e(g_m + e_m) = \frac{(n\,\hbar\,c)}{2}. \tag{4.9}$$

Now compare with our data and use $e_m = e_e$. The fundamental resonance in equation (2.1) corresponds to a theoretical half electron, that is $N = P = 1, K = 0$, and a self-energy $\mu/2$ that we shall ignore. It gives, as per (1 – 3.3):

$$m = X/1 = 8.1451213299073\,\mathrm{keV/c^2}. \tag{4.10}$$

This mass must be compared to μ as it comes from the interaction of the time-currents (not the apparent charges) and then, for an electron, as the product $e^2/4$. The rest of the electron mass (N = P = K = 2) is given by the resonance; then in (4.10) the numbers (N = P = 1) correspond to a hypothetical particle where a current G is interacting with $e/2$ which mass is given by an action corresponding to $Ge/2$.

Now we analyze how action comes as a product of currents, but not energy for which we rely on resonances. In the hypothetical resonance above, it corresponds to the products $e\,G$ and $e^2/4$, where G^2 is absent. It leads to a correspondence between action and energy:

$$\frac{e\,G}{2} \leftrightarrow m; \quad \frac{e^2}{4} \leftrightarrow \mu. \tag{4.11}$$

We divide the two expressions in (4.11) and in light of (4.9) we add $\mu/2$ that we initially ignored; we find:

$$\frac{2\,G}{e} = \frac{m}{\mu} \rightarrow 4\,G + e = 68.4051246306057\,e \approx \frac{e}{2\,\alpha}. \tag{4.12}$$

We *want* to recognize here the modified Dirac condition in (4.5), because the fine structure constant appears linked to the equation parameters.

But the result seems approximate; at first the relative discrepancy (-1.65×10^{-3}) seems acceptable since we analyze a hypothetical particle but we shall see that this numerical value holds precisely.

There is a second aspect related to the Dirac condition which comes from the time-currents model and the apparent electric charges $e/3$ and $2e/3$ going respectively down and up the time; assume their individual self interactions are squared charges. Once again, we can link action and energy:

$$(e/3)^2 + (2e/3)^2 \rightarrow \mu(1/3)^2 + \mu(2/3)^2 = 5\mu/9. \tag{4.13}$$

Now from (4.10):

$$4(m + 5\mu/9)/\mu = 137.032471483434 \approx 1/\alpha \tag{4.14}$$

The relative discrepancy with respect to α is $\approx 2.26\times10^{-5}$. The coincidence can, at first sight, be seen redundant with the equation (4.12) as it is almost identical, but it comes from a different interaction and we shall see now that this value also holds.

4.3 Leptons magnetic moment anomaly

We assumed that the resonances in the previous section "construct" the leptons waves; unlike the classical wave equations the geometrical construction is not unique but lepton-dependent. Thus, even for the electron it seems hardly possible to make an exact link with the Dirac equation which, according to (2.1), should be too general; consequently we go back to de Broglie's thesis which is fully relativistic.

4.3.1 De Broglie wave geometry

In his thesis, de Broglie uses a standing wave, that we will denote the Compton wave and finds a phase wave as a result of the relativistic transformation of the former. The agreement of the stationary wave assumption with the results in Table 1 is straightforward since we get N = P for all leptons.

The change in phase of the de Broglie wave over the first Bohr orbit of a hydrogen atom is $2\,\pi$, while the Compton wavelength change in phase over this orbit is $2\,\pi/\alpha$. Then over any number of Compton wavelengths, we have:

$$\Delta\phi_D = \alpha\,\Delta\phi_C, \tag{4.15}$$

where $\Delta\phi_D$ and $\Delta\phi_C$ are the changes in phase of the de Broglie and Compton waves over any length. On the n^{th} orbit we find:

$$\Delta\phi_D = \frac{\alpha\,\Delta\phi_C}{n}, \tag{4.16}$$

There are n de Broglie wavelengths around the n^{th} Bohr orbit and we get a constant angular differential term α. The same reasoning applies in the case of a nucleus of charge $Z\,e$ and gives the same value. Hence, considering that the de Broglie wave defines the motion of the electron this term is universal in the Bohr model. As a result, and taking into account simultaneously the motion of the electron and the phase velocity of the de Broglie wave going around the proton, the

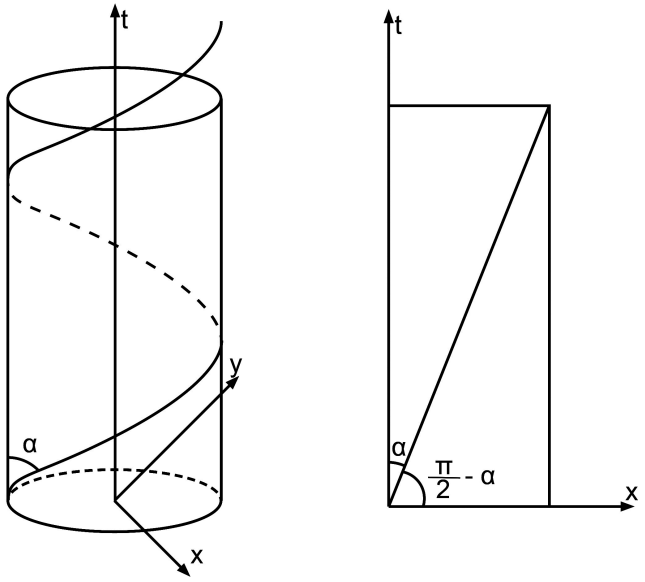

Fig. 1: Left, the electron classical Bohr orbit; right, the same cylinder unfolded (the angle is $\approx \alpha$).

phases of the two waves at any location of the electron classical trajectory are permanently identical.

Assume α is a path length based on integral and geometrical numbers. On the cylinder Figure 1, and using a system of unit where the radius of the cylinder is 1, the length of the unfolded tangent is approximated with $L \approx \sqrt{137^2 + (2\pi)^2}$. Now we know that the electron spin is 1/2, and then the rotation of the resonance is reduced to π when the electron runs one turn; we get the well-known $\sqrt{137^2 + \pi^2} \approx \alpha$.

Consider now the de Broglie wave as a shortcut permanently joining the electron with itself, but one (or n) Compton wavelength later, with an action 1/137/2 (taking again the spin 1/2 into account), it gives:

$$\alpha^{-1} \approx \sqrt{137^2 + \pi^2 - \frac{1}{137} \times \frac{1}{2}}$$

which holds with a relative precision $\approx 3 \times 10^{-8}$. Last, consider that the electron progresses in time, but that its waves are composed of two currents going up and down. If the up-time part of the waves gives a factor 1/2, the down-time part sees the electron with a charge twice lesser since in the case of quarks the down-time and up-time currents manifest fields 1/3 and 2/3 respectively. It must be augmented with a resonance length dependent on the time-velocity of the electron; twice longer for the same reason (charges 1/3, 2/3); finally it gives a factor 1/8 for the down-time part and we get:

$$\alpha^{-1} = \sqrt{137^2 + \pi^2 - \frac{1}{137}\left(\frac{1}{2} + \frac{1}{8}\right)}$$

$$\rightarrow \alpha = 72\,973\,525\,698 \times 10^{-13}$$

which is *exactly* the value of α given in CODATA 2012! Considering precision together with the simplicity of this geometry, it looks pretty much like time-currents exist.

In special relativity, one would consider the so called rapidity of the electron defined as a hyperbolic angle. However, the path length α can also be seen as a simple angle in the Euclidean coordinates (x, y, z, ict) as originally used by Minkowski. Moreover, one must consider this angle universal, and it implies a complimentary angle $\pi/2 - \alpha$. At first the existence of those angles can be checked numerically as it must also correspond to the coincidence (4.3); after appropriate replacements of α^2 by two coefficients corresponding to the two angles α and $(\pi/2 - \alpha)$, the equation (4.3) gives:

$$4\pi\,(m_e - \mu)\,\sin(\alpha)\left[\left(\frac{\pi}{2} - \alpha\right)\sin\left(\frac{\alpha}{\pi/2 - \alpha}\right)\right] = \mu\sqrt{2},$$

which holds with a relative precision of 2.9×10^{-8} instead of 1.25×10^{-5} for (4.3).

4.3.2 Other resonance coefficients and action

When the electron is on the first orbit there is a rotation of the time-current of a hyperbolic angle α which ratio to the space current changes in proportion of the hyperbolic tangent of this angle. As stated, the impact is a phase differential and considering resonances, a simple angle gives $tan(\alpha)$; it runs around the full Bohr orbit and then the instantaneous action term is $tan(\alpha)/2\pi$. The action given by $tan(\alpha)$ is that of a resonance going around the full orbit.

It must cycle on 1/2 quantum; hence the first correction term to the electron magnetic moment anomaly is:

$$a_0^e = \frac{tan(\alpha)}{2\pi} \approx \frac{g-2}{2} \tag{4.17}$$

where we denote a_0 and g the correction and the g-factor respectively. Compare to the first order QED correction by Schwinger [12], the well known $\alpha/2\pi$. The difference comes from a different manner to taking into account relativistic effects. Here it suggests that taking into account together the particle resonances and special relativity in the original Minkowski manner could give an analytic solution. In facts, the difference is that we consider the electron as a 4D gyroscope which axis is bent by velocity. This axis is shown with the orientation of the resonances N, P, K in Figure 2.

Therefore in (2.1) the resonance $N\,P$ corresponds to G^2 in (4.13) while K corresponds to $e^2/4$. The product $N\,P$ makes and "absorbs" the spin and the full space-resonance cycle is then $(N\,P\,-\,2)\,K$ which is a product G^2e^2 while the spin is given by $G\,e$. Action depends on the number of currents C (which, according to the model, is lepton-dependent) while the mass μ is constant; then we divide this coefficient by the number of currents.

We get a spin-dependent coefficient where the spin relates to the interaction of the G-currents and the apparent electric charges — which is logical. It is:

$$E = \sqrt{\frac{NP-2}{C}}K. \tag{4.18}$$

In the direction of time (K in Figure 2), the same reasoning gives NK^2 for a product $e^2/4$. But we get a spin independent coefficient which relates only to the currents and does not need a square root; it is:

$$F = \frac{N\,K^2}{4}. \tag{4.19}$$

The coefficients above are valid for an electron but for the muon and tau the coefficient a_0 corresponding to the time current rotation is not α like in (4.17), it depends on the resonance numbers. The electron is the special case because all resonance numbers are identical and even (N = P = K = 2) and then all phases are identical.

For the muon and the tau, N = P and K are odd and prime with each other, and then the action cycle is $N\,K$. Using (4.18) for an electron, the cycle uses N = K = 2 and its angle should be written $2\,\alpha/2$. Then for a muon and a tau the corresponding coefficient is:

$$\phi = \frac{tan(N\,K\,\alpha/2)}{NK/2}, \qquad a_0 = \frac{\phi}{2\,\pi}. \tag{4.20}$$

The expression mixes angles and resonance and fits with the interaction of current where action is angle-dependent; it will be the geometric form used in this section. We introduce $\alpha/2$ which we now consider as the physical angle of each time-current — it gives α for two currents of opposite directions taken together.

4.3.3 The electron

Now we want to compute the anomaly from the following picture: the electron is seen as a 4D rotation which (in all cases) has the following mathematical property: two orthogonal planes exist which are conserved by the rotation. The identifications are then obvious; the angles in the previous section define the two planes rotations and correspond to the resonances. The rotation is said double since we find distinct angles α and $(\pi/2-\alpha)$. The planes intersect at a single point (a mathematical property of any 4D rotation) where the resonances apply, and it defines the punctual particle — but we do not need to introduce anything material at this place (no particle). The planes intersection point also moves in space and in the direction of time defining a classical trajectory.

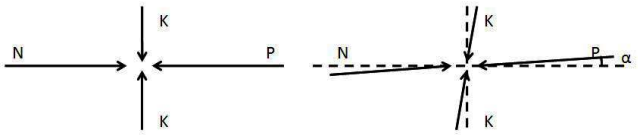

Fig. 2: Resonance geometry on N, P, and K. Left: an electron seen at rest, K on the time axis, N and P in 3-space. Right: an angle $\approx \alpha$ appears as a relativistic shift on the first Bohr orbit where axes are bent by velocity.

One plane is orthogonal to the time axis and hosts the leptons resonances N = P, and K is on the other one which includes the "time translation" of the particle. Finally those two planes are lepton-independent and then their translation and the associated angles define entirely the seemingly anomalous values in (4.8 – 4.10) as they are also lepton-independent. Consequently, the lepton-dependent resonances imply different magnetic moment anomalies. Therefore we can reverse-compute the anomaly from those two quantities. In this way, we define:

From (4.8): $4(X/\mu + 1/2) = \beta_1^{-1} = 136.810249261211$,

From (4.10): $4(X/\mu + 5/9) = \beta_2^{-1} = 137.032471483434$.

The Dirac equation gives $g = 2$ and it is known that the correction is entirely related to relativistic shifts. The quantities above correspond to distinct interactions and then distinct types of charges; hence the correction is a product $a_T = a_0\,a_1\,a_2$ where a_0 is geometrical and corresponds to the angle α in (4.17) or ϕ in (4.20), a_1 to the action of the apparent electric charges (4.10), and a_2 to the action of (magnetic) currents (4.8).

Since β_1 and β_2 are deduced from the leptons masses, they are related to the tangent of some angles part of the resonance geometry (in the same manner as $tan(\alpha)/2\pi$). The anomaly is angular and differential and then a_1 and a_2 must be computed as ratios involving α and the arctangents of some angles involving respectively β_2 or β_1, and resonance numbers. The electron correction term a_1^e is then given by an expression of following form:

$$\frac{\tan(\alpha)Y}{\tan^{-1}(\beta_2 Y)} \to a_1^e.$$

It links an action given by the angle α and another one given by β_2 and the anomaly relates to their ratio. Now β_2 relates to the apparent electric charges giving the spin; then $Y = E$ as defined in (4.18). The angle $\alpha/2$ also impacts the coefficient and subtracts from K.

Then we write:

$$E \to \sqrt{\frac{NP-2}{C}\left(K+\frac{\alpha}{2}\right)} \tag{4.21}$$

$$a_1^e = \frac{\tan(\alpha)\,\sqrt{2+\alpha/2}}{\tan^{-1}\left(\beta_2\,\sqrt{2+\alpha/2}\right)} \tag{4.21.1}$$

Now β_1 comes from the time-currents of the electron; we must make a similar reasoning involving F defined in (4.19).

Naturally, this correction will be similar in form to the equation above. The logic is:

- The first order effect is null; it is second order where the cross-products cancel.
- The angle must be α instead of $\alpha/2$ since the two angles $\alpha/2$ on the axis of K sum up.

It gives, for an electron:

$$a_2^e = \frac{\tan(\alpha) F\left(1-\alpha^2\right)}{\tan^{-1}\left(\beta_1 F\left(1-\alpha^2\right)\right)}, \tag{4.22}$$

$$a_2^e = \frac{\tan(\alpha)\left(2-2\alpha^2\right)}{\tan^{-1}\left(\beta_1\left(2-2\alpha^2\right)\right)}. \tag{4.22.1}$$

Note that in the equations (4.21 – 4.22) the angle $\alpha/2$ affects K and $-\alpha^2$ affects K^2; it is the same geometry where only K is impacted. Now from (4.17 – 4.21.1 – 4.22.1) and using the value of α in CODATA 1014 we find:

$$g_T^e/2 = 1 + a_0\, a_1\, a_2 = 1.00115965218091. \tag{4.23}$$

The values of X and μ in (3.1) were tuned to fit with CODATA 2014 which gives:

$$g^e/2 = 1.00115965218091\,(26). \tag{4.24}$$

The relative error on g_T^e in (4.23) with respect to (4.24) is less than 10^{-14}, but it can be down to $\approx 10^{-8} - 10^{-9}$ without ad-hoc tuning and keeping all leptons masses within uncertainty — the result would still be very significant.

4.3.4 The muon and tau

We get the equations needed to compute the muon anomaly in the same manner as for the electron but using (4.20) and including in (4.21) the four currents given by the model, and the resonance numbers in Table 1. We get:

$$g_T^\mu/2 = 1.00116592081. \tag{4.25}$$

The CODATA 2014 experimental value is:

$$g^\mu/2 = 1.00116592089\,(63). \tag{4.26}$$

The result is well within experimental uncertainty and independent of the adjustments of (3.1) since the precision is in the range 10^{-9}. The SM prediction disagrees with a $2-4\sigma$ discrepancy. Typically:

$$a^\mu_{SM} - a^\mu_{experiment} = (2.8 \pm 0.8) \times 10^{-9}. \tag{4.27}$$

The very short lifetime of the tau makes impossible at present to measure its $(g-2)$. The SM prediction is:

$$g^\tau_{SM}/2 = 1.00117721\,(5). \tag{4.28}$$

Using the tau resonances in Table 1 we get:

$$g_T^\tau/2 = 1.00125789. \tag{4.29}$$

But on the other hand, in the tau resonance, N = P = 9 is not a prime number, it is a square and then, perhaps, we should use 3 instead of 9 in the equations to compute its anomaly (we find a second reason later). It gives:

$$g_T^\tau/2 = 1.00117037, \tag{4.30}$$

where the difference with the SM prediction is more coherent with that of muons.

4.4 The fine structure constant

We made a first calculus of α as a simple path length. Now we shall first show that the shortcuts in this path length, namely 1/2 and 1/8, also defines the leptons resonances, and then find an immediate origin to the number 137.

4.4.1 A second view on leptons resonances

Our analysis of the resonances in Table 1 fits with the supposed geometry, and complimentary angles α and $(\pi/2-\alpha)$. It is a quasi-symmetrical picture that suggests the existence of a second view on the leptons resonances agreeing with the equation (2.1). In this equation we use three resonance terms (N, P, and K), but the rotation is in 4 dimensions; then the resonance terms correspond to one rotation plane used completely (N, and P), while K lives in the other plane but we only use an axis (not the full plane). The second view should split oppositely; it cannot hold with $N = P$ but it must with $P = K$ because of phase coherence. Then using angular ratios, we should have a different mass: $\mu' \approx \mu\,\pi/2 \approx 380$ eV/c^2. Starting with this value, imposing P = K, and using the equation (2.1), an empirical fit to the same decimal as shown in Table 1 gives Table 5 and the coefficients in (4.31).

Table 5: Second view on electron, muon, tau in MeV/c^2.

–	P=K	N	Computed	Measured
e	2	2	0.510 998 9461	0.510 998 9461(31)
μ	3	8	105.658 3752	105.658 3745(24)
τ	4	16	1 776.84	1 776.82(16)

$$\begin{aligned} \mu' &= 385.6750521055\,\text{eV/c}^2, \\ D' &= 0.0002255984538, \\ X' &= 8.02160795579\,\text{keV/c}^2. \end{aligned} \tag{4.31}$$

$P = K$ is verified, and we can estimate:

$$\mu' = \mu\left(\frac{\pi}{2} + \frac{\pi}{137} + \left(\frac{2\pi}{137}\right)^2\right),$$

which was used to compute (4.31); it uses 1/137 and no simple fit was found with α.

The remarkable point in Table 5 is that we find for N the numbers 2 and 8, and their product 16 for the tau. Those numbers show that, in the EM field, the resonance is tachyonic and the shortcuts can ring independently or in a combined manner. The product 16 also justifies our doubts for the tau $(g-2)$ in (4.29).

4.4.2 Alpha and 137

Following the first equation giving α, assuming time-currents exist and correspond to $e/2 \to 1/274$ we find an empirical fit compatible with CODATA 2014:

$$\alpha^{-1} = \sqrt{137^2 + \pi^2 + \frac{1}{274^2} - \frac{1}{137}\left(\frac{1}{2} + \frac{1}{8}\right)}$$

$$\to \alpha = 72\,973\,525\,672 \times 10^{-13}, \qquad (4.32)$$

where the difference with CODATA 2014 is about half the standard deviation:

$$\alpha_{CODATA\,2014} = 72\,973\,525\,664\,(17) \times 10^{-13}.$$

But now, why 137? A straightforward calculus gives a possible origin; taking all integral N and P from all tables, we get a seemingly absurd suite of numbers that sums to:

$$\Sigma_{NP} = 2 + 3 + 4 + 5 + 7 + 8 + 9 + 12 + 14 + 16 + 19 + 38$$

$$= 137. \qquad (4.33)$$

Is that a coincidence, or rather the signature of a discrete wave packet? If one thinks of exponentiation, each term of the sum corresponds to a different piece of the phase of a unique signal which includes all symmetries and all the manners they combine, interact and condense (or ring). Since N and P are space currents, Σ_{NP} defines a universal oscillator. With respect to field theory, it is straightforward that such a wave includes or represents all virtual particles fields.

A complimentary result on $K \to 274$ seems doubtful; however, taking 266 from bosons instead of (-2, -7, -19), and the distinct values of K from leptons and quarks, we notice:

$$\Sigma_K = (2 \times 7 \times 19) + 2 + 3 + 4 + 5 - 6 = 274. \qquad (4.34)$$

The interpretation is less obvious and the link with known theory is nil, because this quantity addresses the effect in space of vibrations or rotations along the time axis and their participation to particles mass and interactions; there is no such concept in known theory.

In any case, those relations are complimentary to each other and provide with numerical coherence linked to the concepts developed before.

4.4.3 Splitting D_e and D'

Now, α is a 4D path length as seen in 3+1D, then the couplings D' (4.31) and D_e (3.1) should have a similar form but in a complimentary manner with respect to the resonance terms; hence they should also be expressed with similar expressions but using 3, 7, and 19 (the resonances of quarks) and $\Sigma_K = 274$; we find the following empirical fit which terms show an obvious symmetry:

$$D_e^{-1} = \sqrt{((7-3) \times (274+19))^2 + 7\pi^2 - \frac{19\pi}{19-1}}, \qquad (4.35.1)$$

$$D'^{-1} = \sqrt{((19-3)(274+3))^2 + 2^2 \times 3 \times 7\pi^2 - \frac{3}{3-1}}. \qquad (4.35.2)$$

Those expressions were used to compute the values in (3.1 – 4.31) and then all masses.

Several aspects are remarkable in those expressions:

- We notice that 274 + 3 = 277 and 274 + 19 = 293 are also prime numbers; hence those are not reducible. Their difference is 16 which is also $(7-3)^2$ in D_e and $(19-3)$ in D'.
- The rotation term $7\pi^2$ in (4.36.1) is a perfect fit with the μ and τ resonances (5 = 7 – 2, and 9 = 7 + 2), where 7 was inferred a rotation.
- D' includes a factor 2, which can be inserted in K in Table 5, but not in P; then P and K act on the time and magnetic moment axis respectively and it must be identical to the classical g-factor = 2. This is necessary since Table 5 is in the symmetry of QED.

Importantly, the expressions above are obtained by simple divisions based on the initial empirical fit of the D_e and D'. The left term is the closest square to the empirical value of D^{-2} from which it is subtracted; the middle integral term is the division of the rest by π^2 that gives a small residual term. Then we search to express all terms with integral numbers — preferably those we expect.

5 Gravitation, the keystone

The mass equation and the time-current model are coherent with Cramer's transactional interpretation of quantum mechanics which fills the gap of non-locality (the true signature of quantum physics) but without spooky action in 3-space.

Since the reasoning to the mass equation (thru N and P) and Cramer's interpretation are relevant in absorber theory and uses a pressure field, gravity must be analyzed in a shielding manner using Wheeler-Feynman equations [13, 14]; in this way, it was shown compatible with gravitation in a recent paper [6]. It does not require the existence of dark matter to explain the observations at the origin of this hypothesis and it also explains the cosmos energy densities (visible, dark, and visible + dark). In this section, we shall not restate the piece of theory in [6] but only the logic and main results.

The absorber free energy equivalent mass M_A is given by symmetry of the absorber process in gravitation; we first write the energetic part of the Schwarzshild metric:

$$c^2\,d\tau^2 = (c^2 - 2\,G\,m/r)\,dt^2 - \frac{c^2}{c^2 - 2\,G\,m/r}(dx^2 + dy^2 + dz^2)$$

Then, in the spirit of absorber theory, we symmetrize the equation in geometry and mass terms:

$$\frac{2\,G\,m}{r} = \frac{m\,R_u}{M_A\,r} \rightarrow \frac{R_U}{2\,M_A} = \frac{G}{c^2}, \tag{5.1}$$

$$\rightarrow M_A = \frac{R_U\,c^2}{2\,G} = \frac{P_p\,T}{2\,c^2} = 9.790 \times 10^{52}\ \text{kg} \tag{5.2}$$

where $R_U = cT$, $T = 1/H$ is the age of the event horizon while H is the Hubble factor and P_p is the Planck power.

Concerning visible energies $M_V c^2$, the ratio M_V/M_A is a geometrical constant. This constant links a 4-volume and a linear interaction in 3-space; the surface of a 4-sphere is $2\,\pi^2\,R^3$, and then the factor 2 in (5.2) becomes $4\,\pi^2$ in $3 + 1D$ where visible energies interact thru the light cone. It gives:

$$\frac{M_A}{M_V} = 2\pi^2 \rightarrow M_V = 4.453 \times 10^{51}\ \text{kg}. \tag{5.3}$$

Summing, we get the total energy M_U of the visible universe:

$$M_U = M_A + M_V = 9.236 \times 10^{52}\ \text{kg}. \tag{5.4}$$

It gives to a total density $\rho = 9.91 \times 10^{-27}$ kg/m^3 and the visible energy (5.3) is 4.82% of the total. The benchmark at this time is the Planck mission results [20] which gives $\rho = 9.90\,(6) \times 10^{-27}$ kg/m^3 and 4.86 (10)% of visible matter. Hence according to the standard model of cosmology we get valid quantities. The equation (5.1) also means that the rate of dark energy creation (M_A) since the initial bang is constant and half the non-reduced Planck power: the universe energy is identical to its expansion and we do not find a big bang but a permanent process. Next, using the Wheeler-Feynman equations or Newtonian gravity this creation gives an acceleration excess up to $H\,c$ at the galaxy borders, meaning the absence of dark matter.

But now what is the relation with our analysis of mass? According to (4.34 – 4.35), the numbers 137 and 266 address space and time respectively. They interfere at the point of origin which is visible thru the solid angle 4π, and we should find there the reduced Planck mass giving the Planck power:

$$M_p = \sqrt{\frac{hc}{G}} \times \frac{1}{4\pi} = 2.43536\,(6) \times 10^{18}\ \text{GeV/c}^2.$$

Using the mass equation (2.1) with the parameters in (3.1) and taking $N = P = 137^2$, and $K = +1/266^2$ gives a mass:

$$M = 2.464 \times 10^{18}\ \text{GeV/c}^2,$$

which is very close to Mp.

Looking at (4.36.1), we find $7\pi^2$ in the expression of D_e while 19 has a role similar to 7 in the case of quarks (N, Table 2) and bosons (K, and N = P = 19 – 7 in Table 4); then in order to symmetrize the equation we take:

$$N = P = 137^2 - 19\,\pi^2;\ K = +1/266^2,$$

$$M = 2.43526 \times 10^{18}\ \text{GeV/c}^2.$$

Finally, the next two decimals are given by addition of $\approx 2/3$ to $N = P$; a small empirical term which is expected as it makes this expression homogeneous to coupling:

$$N = P = 137^2 - 19\,\pi^2 + 2/3;\ K = +1/266^2, \tag{5.5}$$

$$M = 2.43536 \times 10^{18}\ \text{GeV/c}^2. \tag{5.6}$$

Since $1/(N\,P) < K\,D$ this resonance is not permitted in 3+1D.

Considering that we now discuss reconciliation of quantum theory and general relativity through a common origin this result is keystone on top of the study. It shows that the same field also leads gravitation.

Here we can define a unit-less quantum gravitational coupling constant which reads:

$$\alpha_G = \frac{X^2}{M_p^2} = \left(\frac{1}{(137^2 - 19\pi^2 + 2/3)^2} + \frac{D_e}{266^2}\right)^6, \tag{5.7}$$

$$\alpha_G = 1.1186 \times 10^{-47},$$

where we see that the rest of quantization lives in and from a single oscillator defining gravity; it is "below" quantum theory and it does not need the existence of a graviton particle.

Unlike the classical definitions of α_G, since X is universal and represents the pressure field, (5.7) is unique and does not depend on an arbitrary choice of mass.

But now the ratio of the electron mass to the Planck mass is constant, which seems a contradiction with (5.1). On the other hand, the observable cosmos has constant atomic physic and chemistry and then its laws use relative constants varying in time and not absolute ones. Thus, only unit-less quantities are constant; since G is used with constant masses in classical theories, then hc and G vary together in the same manner as $\alpha = e^2/\hbar c$ is constant.

Therefore, here is the big picture, the minimal interpretation of all results in this paper (no doubt it can be made more complex and elegant):

- A Planck particle exists at the origin; it emits a wave of Planck length and time. This resonance exists in 4 dimensions, it is not energy but its wave defines the quantum of action.
- This wave interacts thru the light cone (and gives 137 in α), and thru a radial line (giving 266 and 274). In a symmetric absorber concept, it means that the universe and its origin are quantizing each other.

- The emission is constant and corresponds to the Planck power; it builds M_A, and the visible energies field M_V is the absorber. It creates a deficit which is gravitation (see the absorber equations in [6]).
- For complete quantization thru time-currents, 137 is the sum of all resonances in space, and 266 is the product of the bottom space-type resonances, radial or circular (2, 7, 19).
- Increasing masses and the constancy of $e^2/\hbar c$ and hc/G are equivalent to and interpreted as time dilation. It denotes the emergence of the observable time in a frame where it does not exist. The observable time is seen as a radial progression in 4D space.

6 Discussion

Firstly, what have we been discussing all along? Essentially, the reasoning to the mass equations (2.1 – 2.2) is based on the existence of a stationary wave in a universe where:

a) everything propagates,

b) mass and charges do not have a proper existence,

c) the field is self-quantizing, and consequently a unique field and mechanism exists.

We end-up with a wide picture where all (free) parameters of the SM related to energy are self-quantizing geometry of movement (at this stage, and taking all results above, only the SM parameters expressed as phases or angles are not computed); the same is valid in gravitation and cosmology. *Hence we discuss the very nature of energy*, of its forms and formation on top of a unique field; something looking like the natural reductionist path of science.

Secondly, what does it means with respect to the standard theories? In its present form QFT neither considers *definite* rotations nor signals going up and down the time. Therefore no true comparison with our results is possible. Still, we find a number of connections like coupling constants and other aspects which will be discussed in the next paragraph.

In cosmology and using general relativity, a permanent energy creation is not even envisioned. Still, energy conservation comes from time-translation invariance and Neuter's theorem; but we know that the background (R_U) increases and then there is no mathematical reason that energy is conserved in cosmology.

The third point to discuss is the possibility of a different universe (a fashion question). But it seems unlikely because, as shown before, all resonances decay from $2, 3, 7 = 2^3 - 1$, and $19 = 3^3 - 2^3$; then probably only 2 and 3. It leads to conjecturing further the role of symmetry in the mass equation; essentially how do we get 2, 3, 7 and 19, and what is the limiting factor if any? Now let us reason on this aspect.

In the mass equations, the resonances N and P should come straight from the equation geometry and group theory. We shall use 1, 2, 3 to denote U(1), SU(2) and SU(3) respectively and discuss field polarization in the resonance equation. With field polarization p we mean dipoles or tri-poles where summing p charges makes a neutral. In the following, one must just keep in mind that $U(1) \subset SU(2) \subset SU(3)$.

- At the core of a particle resonance, time currents give a charge Q constant; its polarity is p (in 2, 3). In any sphere centered on Q the sum of charges is Q. Then except for Q, the total charge separation in a scale-independent 3-sphere depends on a cube, say n^3 (since the resonance radius is arbitrary) and it is neutral.
- In the radial case, with resonance P, on each layer of the resonance the radial action is layer independent, then the radial coefficient of polarization in $1/n^2$ for each layer, with $1 \leqslant n \leqslant P$; then $P = n$. The polarity of Q is p and defines the interaction of the particle which is also radial, then on the radial path $n = p = P$. Here P defines simultaneously a radial exchange of action and polarity (the symmetry). This is immediately verified for quarks ($P = 3$), and for leptons if we decompose P as shown before.
- On a circular path, a resonance N gives N circular sectors with identical action and action coefficients. Then $N = n^3$ on this path. Since this number does not define the radial interaction of the particle, any subgroup of p is acceptable, then $1 \leqslant n \leqslant p$.

We get the following suites of numbers:

- On a radial path the polarity is p, and $P = p = 2$ or 3;
- On a circular path the polarity is n with $1 \leqslant n \leqslant p \rightarrow N = n^3$; limited to $2 \rightarrow 8, 3 \rightarrow 27$.

But the latter is a rotation, not a resonance as needed, and we need to complete the reasoning.

With geometry and currents (and nothing else), the logical manner is to combine symmetries. Say in the resonance volume we have two symmetries at work; a structural point of equilibrium needs a transformation. Therefore, on the circular path a resonance is seen as a transformer in $n \leq p$ and the subgroups of n, where coefficients are the same for n and its subgroup. Hence, on circular paths we get cubes differences 7 and 19; those come like transformer of charges or currents between a group and its sub-group. That is to say that the field polarization $n \rightarrow n^3$ is always balanced by $(n-1) \rightarrow (n-1)^3$. Importantly, there is nothing in this reasoning preventing more complex oscillators, for instance $19 - 7 = 12$.

This discussion leads to introduce U(1) which is a very special case; since $1^3 - 0^3 = 1$ it seems to be a massless field with any oscillator; the same reasoning on 0 suggests a continuous current — an amplitude according to which masses and then the observable lengths and the rate of time vary in reverse proportions.

Now why only 0, 1, 2, and 3? Within the logic above, the first mathematical explanation is Hurwitz theorem [10]. Consider two charges or currents x and y, we may need to

compute the impact of y on the self-interaction of x; it is equal to the action on x of the interaction of x and y (conservation and symmetry), then:

$$(x\,x)\,y = x\,(x\,y).$$

This is the definition of alternative algebra; according to Hurwitz theorem [10], only four exist which are R real numbers, C complex numbers ~ U(1), H quaternions ~ SU(2), and O octonions ~ SU(3). One can consider this as limiting either the symmetry spectrum, or just our ability to model with particles and charges — or both.

A peculiar case arises with $x\,X = 1$ or unitary; the impact of $x\,X$ on any other quantity of the same group does not change its amplitude. Then $x\,X$ addresses structural conservation and we find simultaneously 137, $1/137$, and 274, $1/274$ in the expression of couplings.

In this way those quantities are related to the monopole as quantized rotational 4D paths, like α and D_e, where only couplings can be measured in 3+1D as seemingly arbitrary real numbers.

7 Conclusions

The breakthrough to wave equations was the assumption of a stationary wave pervading all space. But how can such a wave exist in relativity without a mass of its own? How could it be distinct from the mass of the particle or system it describes? Then how could it be distinct from gravitation?

Those naive but unsolved questions are almost a century old as they address the nature of the wave, wave-particle duality, the completeness of quantum mechanics, and the physical link between gravitation and quantum physics.

The novelty here is that those questions are justified by the existence of a solution to the free parameters problem, including and linking particles physics, gravitation, and cosmology, not only by conceptual disagreements or theoretical incompatibilities.

As stated in introduction, we do not solve any equation; the existence of a solution is first seen when the mass equation is fit to phenomenology, and then extended to couplings. We find logical coherence, a reductionist concept and fantastic precision. Of course it does not look like the usual manner in modern physics where theory and principles reign; but, considering the difficulty of solving this problem from theory, it might be the only practicable way — at least at present time.

As a matter of conclusion, it looks as though the solution shown here can be found only as a whole and provided that we do not build on existing concepts (and maybe even principles); but one must first recognize the existence of a problem together with its ramifications. This situation is fantastic and terrible; if *that* solution exists, physics could remain stuck endlessly in its present conceptual state *because of this conceptual state*: whatever new particles discovered in collision machines modeled with ad-hoc SM extensions, its framework may never be contradicted by experiment.

8 Addendum

As for the 750 GeV resonance possibly detected at CERN [21], since it decays to two photons we assume the same equation and parameters as the H^0 and only K can be fit; it gives $K = -133/2$ which is immediately remarkable. However, since K is not integral the width must be reconsidered, logically to $\Delta K = 1/4$, giving from (2.2 – 3.6 – 3.9.3):

$$N = P = 12, K = -133/2 \rightarrow m \approx 744.9\ \mathrm{GeV/c^2},$$

$$\Delta K = 1/4 \rightarrow \Gamma \approx 9.6\ \mathrm{GeV/c^2}.$$

Using (3.9.2) instead of (3.9.3) adds +3 $\mathrm{GeV/c^2}$ to the mass. The other candidate with $\Delta K \approx 1$ gives $\Gamma \approx 40\ \mathrm{GeV}/c^2$.

At this scale, the equation (2.2) is very sensitive to D and the model in time-currents must be identical to the H^0 otherwise the computed mass is far from the estimate. It would be very similar, but it leads to remark that there are two manners to put four distinct charges at the corners of a tetrahedron; there may be a chiral difference with the H^0, justifying distinct masses and a probable impact on the particle decays.

Last, the number 133/2 verifies (2.3) like 7 and 19, but with $P = 1$ instead of $P = 3$, since $133\,\pi/2 \approx 209/1.0004$. It is even doubly remarkable since 209 is multiple of 19.

Hence the existence of this particle, if confirmed, should not change the values of Σ_{NP} and Σ_K; it fits well and naturally with the logic and results in this paper.

Submitted on April 3, 2016 / Accepted on April 30, 2016

References

1. Bethe H.A. The electromagnetic shift of energy levels. *Phys. Rev.*, 1947, v. 72, 339.
2. De Broglie L. Recherches sur la théorie des quanta. *Annales de Physique* — 10e sérié — Tome III — Janvier-Février 1925.
3. Consiglio J. Below the Standard Model? *Applied Physics Research*, v. 6(2), 19–27, 2014.
4. Consiglio J. On the Field Below SM, Analysis and Predictions *Applied Physics Research*, v. 6(5), 1–17, 2014.
5. Consiglio J. Hacking the Fine Structure Constant in Leptons Geometry *Applied Physics Research*, v. 7(6), 30–46, 2015.
6. Consiglio J. On the Absorber in Gravitation. *Progress in Physics*, v. 12(1), 20–25, 2016.
7. Cramer J. The transactional interpretation of quantum mechanics. *Rev. Mod. Phys.*, 1986, v. 58(3).
8. Dirac P.A.M. Quantized singularities in the Electromagnetic Field, *Proc. Roy. Soc. A*, 1931, v. 133, 60.
9. Dirac P.A.M. (1948). The Theory of Magnetic Poles. *Phys. Rev.*, 1948, v. 74, 817.
10. Dickson L.E. (1919). On Quaternions and Their Generalization and the History of the Eight Square Theorem. *Ann. Math.*, 1919, v. 20, 155.
11. Poincaré H. Sur la dynamique de l'électron. *Rendiconti del Circolo Matematico di Palermo*, 1903, v. 21, 129–176.
12. Schwinger J. On Quantum-Electrodynamics and the Magnetic Moment of the Electron. *Physical Review*, 1948, v. 73(4), 416.
13. Wheeler J.A., Feynman R.P. Interaction with the Absorber as the Mechanism of Radiation. *Reviews of Modern Physics*, 1945, v. 17(2–3), 157–161.

14. Wheeler J.A., Feynman R.P. Classical Electrodynamics in Terms of Direct Interparticle Action. *Reviews of Modern Physics*, 1949, v. 21(3), 425–433.

15. ATLAS collaboration. Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *(*Phys. Rev. D), 2014, v. 90(5), 052004.

16. CMS collaboration. Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model (Technical report). 2014, CMS-PAS-HIG-14-009. CERN. https://cds.cern.ch/record/1728249?ln=en

17. ATLAS and CMS Collaborations. Combined Measurement of the Higgs Boson Mass in pp Collisions at sqrt(s)=7 and 8 TeV with the ATLAS and CMS Experiments. 2015. arXiv: 1503.07589.

18. The ATLAS, CDF, CMS and D0 Collaborations. First combination of Tevatron and LHC measurements of the top-quark mass. 2014. arXiv, 1403.4427

19. The CMS Collaboration. Measurement of the top-quark mass in top-quark pair events with lepton+jets final states in pp collisions at sqrt(s)=8 TeV. 2014. CMS Physics Analysis Summary TOP-14-001. https://cds.cern.ch/record/1690093?ln=en

20. The Planck Collaboration. Planck 2015 results. I. Overview of products and scientific results. arXiv: 1502.01582

21. ATLAS note, ATLAS-CONF-2015-081, Search for resonances decaying to photon pairs in 3.2 fb-1 of pp collisions at sqrt(s) = 13 TeV with the ATLAS detector. CMS note, CMS PAS EXO-15-004 Search for new physics in high mass diphoton events in proton-proton collisions at 13 TeV.

On the Nature of Ball Lightning

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

The author proposes a model of ball lightning based on a mechanistic interpretation of John Wheeler's ideas. It is assumed that ball lightning is a quasi-particle that has the Planck mass and consists of closed contours, which in turn are based on the magnetic and gravity force balance. These contours are hard packed in a small volume of ball lightning, forming a multilayer capacitor containing a substantial charge and electrostatic energy. This paper provides calculations of characteristic parameters of ball lightning, which are well consistent with its phenomenology.

1 Introduction

There are many theories about ball lightning. However, the nature of this mysterious phenomenon remains unclear. This paper proposes a model of ball lightning based on a mechanistic interpretation of John Wheeler's concept. Previously, such an approach has been successfully applied to construct both micro-world and space models (see [1–4], etc.).

To some extent, the model proposed is similar to the quantum model by Geert Dijkhuis, Professor at Eindhoven University of Technology and Secretary of the International Committee on Ball Lightning. His model suggests that ball lightning is a macroscopic quantum object. Earlier, a similar hypothesis was proposed by Boris Ignatov [4]. Nevertheless, it must be noted that there is no complete understanding of the nature of such objects. It is assumed that a quasiparticle cannot carry a substance. It only carries energy, pulse, and momentum, while the electrons inside such an object are completely coherent and make up a single wave function. Energy of such a quasiparticle is gradually dissipating, in the visible range in particular. Therefore, ball lightning can be observed as an optical object.

2 Presuppositions

Recall that (according to Wheeler) there are original primary elements of space and matter, which have different names — *wormholes, appendices, current tubes, threads or force lines of a field*. If they are real objects and not just mathematical abstractions, in physical terms, they must be some kind of vortex structures resting on the phase boundary (surface). In particular, Wheeler treats charges as singular points on the surface source-drain connected by current tubes in an additional dimension forming a closed *contour*.

Paper [1] shows that, from a purely mechanistic point of view, the *charge* is proportional to its momentum about the contour of the vortical tube and reflects the extent of non-equilibrium of physical vacuum; *spin* is proportional to the angular momentum relative to the longitudinal axis of the contour, respectively; and the *magnetic interaction* between the conductors is similar to forces existing between the current tubes. It is customary that a single element of such a tube is an element with the size of a classical electron radius r_e and its mass m_e.

The model of ball lightning was only built using the ratio of fundamental interactions as in the above-mentioned papers of the author. The mechanistic interpretation of Wheeler's ideas makes it possible to record formulae for electric and magnetic forces in a *Coulomb-free* form, where the charge is replaced with the ultimate electron momentum. In this case, the electric and magnetic constants, ε_0 and μ_0, are as follows:

$$\varepsilon_0 = \frac{m_e}{r_e} = 3.33 \times 10^{-16}\ \text{kg/m}, \tag{1}$$

$$\mu_0 = \frac{1}{\varepsilon_0 c^2} = 0.0344\ \text{N}^{-1}, \tag{2}$$

where the electric constant becomes linear density of the vortical tube, and the reciprocal of the magnetic constant is the centrifugal force produced by rotation of the vortical tube element with m_e mass with the velocity of light c along r_e radius. This value is also equivalent to the force existing between two elementary charges at the given radius.

For the purpose of mutual comparison of interactions, formulae for the electric, magnetic, gravity, and inertial forces are written in a dimensionless form with a single dimension factor of force $1/\mu_0$. With (1) and (2) in mind, we have the following:

$$F_e = \frac{1}{\mu_0}\left(\frac{r_e}{r_0}\right)^2 z_{e_1} z_{e_2}\,, \tag{3}$$

$$F_m = \frac{1}{\mu_0}\left(\frac{l}{2\pi r_0}\right)\left(\frac{r_e}{c \times [\text{sec}]}\right)^2 z_{e_1} z_{e_2}\,, \tag{4}$$

$$F_g = \frac{1}{\mu_0}\frac{1}{f}\left(\frac{r_e}{r_0}\right)^2 z_{g_1} z_{g_2}\,, \tag{5}$$

$$F_i = \frac{1}{\mu_0}\frac{r_e}{r_0}\left(\frac{V_0}{c}\right)^2 z_g\,, \tag{6}$$

where V_0, r_0, l, z_e, z_g, and f stand for circumferential velocity, circumferential radius or distance between vortical tubes, length of the vortex tube (thread) or contour, relative values of the charge and mass of the electron charge and mass, and the

electric-gravity force ratio, respectively, with the latter having the following formula with the same designations:

$$f = \frac{c^2}{\varepsilon_0 \gamma} = 4.16 \times 10^{42}, \tag{7}$$

where γ is the gravitational constant.

3 Calculation of characteristic parameters of ball lightning

Ball lightning often originates from streak one. Imagine streak lightning as a bundle of vortex threads, which (under certain conditions) form vortical current tubes. The latter, in turn, are closed into contours. It is obvious that there must be balances of some pairs of interactions for ball lightning to exist. They are the following:

1. Ball lightning mass M satisfies the condition of equality of *electric and gravity forces*, so with unit charges we have the following:

$$M = f^{1/2} m_e = 1.86 \times 10^{-9} \text{ kg}, \tag{8}$$

which is in agreement with the Plank mass by order of magnitude.

2. Closed contour branches with opposed currents satisfy the balance of *magnetic and gravity forces* resulting in a linear geometric mean dimension of the contour:

$$l_k = (r_0\, l)^{1/2} = \left(\frac{z_{g_1} z_{g_2}}{z_{e_1} z_{e_2}}\right)^{1/2} \left(\frac{2\pi}{f}\right)^{1/2} c \times [\text{sec}], \tag{9}$$

where the ratio of the product $\varepsilon = (z_{g_1} z_{g_2})/(z_{e_1} z_{e_2})$ is an *evolutionary parameter* that characterizes the state of the environment and its changes as the mass carriers dominate over electric ones, and in fact shows the distinction between material medium and vacuum. Hereinafter, we shall take it as being close to the unit in our case, while $l_k = 3.68 \times 10^{-13}$ m, and the vortical tube's mass is

$$m_k = \varepsilon_0 l_k = 1.19 \times 10^{-28} \text{ kg}. \tag{10}$$

In addition, if we express vortical tubes' masses in (9) — $z_g m_e$ as $\varepsilon_0 l$, then we shall get the following relation between the contour axes for unit charges with (7) in mind:

$$\frac{r_0}{l} = 2\pi \rho_e \gamma \times [c^2] = 17070 \approx a^2, \tag{11}$$

where ρ_e is the electron density equal to $m_e / r_e^3 = 4.071 \times 10^{13}$ kg/m^3; and a is the reverse fine structure constant equal to 137.036. Thus, the individual contour is most likely to have axes equal to the size of an electron r_e and Bohr atom $r_e a^2$.

3. Vortical tubes of a contour consist of a number of unidirectional parallel individual vortex threads spinning about the longitudinal axis of the contour with circumferential velocity V_{0i}. Their stability is ensured by the balance of *magnetic and inertial forces*, which give rise to the following formula:

$$V_{0i} = \left(\frac{z_{e_1} z_{e_2}}{z_g}\right)^{1/2} \left(\frac{r_e l}{2\pi}\right)^{1/2} \frac{1}{[\text{sec}]}. \tag{12}$$

Individual vortex filaments having length l and mass carriers in the number of $z_g = l/r_e$ are spinning about the longitudinal axis along an indefinite radius. In the case of unit charges, we have the following minimum circumferential velocity about the longitudinal axis:

$$V_{0i} = \frac{r_e}{(2\pi)^{1/2} \times [c]} = 1.124 \times 10^{-15} \text{ m/sec}. \tag{13}$$

The total number of contours (and the same of unit charges, respectively) may be as follows:

$$z = \frac{M}{m_k} = 1.56 \times 10^{19}. \tag{14}$$

The way these contours are packed in the volume of ball lightning is unclear. Possibly, a contour may be one-dimensional with the total length of $z \times l_k$. It can be expected that with transformation into the more energetically favourable structure the contour (folding repeatedly) forms a large number of loops or cells, which are enclosed in a spherical volume with a bright centre (nucleus). In both cases, with the elements being the most densely packed in the volume, the reduced minimum linear dimension of the outer spherical surface will be as follows:

$$l_{min} = z^{1/2} l_k = 0.00145 \text{ m}, \tag{15}$$

However, if we consider the ratio in (11) and take one of the axes of the Bohr radius instead of l_k for individual contours, then we can estimate the maximum size of such a sphere as $l_{max} \approx 0.00145 \times a \approx 0.2$ m.

Let us calculate the rest limit parameters of ball lightning — energy, charge, electric potential of streak lightning required for generation of ball lightning, and its ultimate density:

$$E_{lim} = Mc^2 = 1.67 \times 10^8 \text{ J}, \tag{16}$$

$$q_{lim} = z e_0 = 2.50 \text{ K}, \tag{17}$$

$$U_{lim} = \frac{E_{lim}}{q_{lim}} = 6.68 \times 10^7 \text{ V}. \tag{18}$$

The density will be calculated taking l_{min} as the sphere's diameter:

$$\rho_{lim} = \frac{M}{\frac{4}{3}\pi \left(\frac{l_{min}}{2}\right)^3} = 1.17 \text{ kg/m}^3, \tag{19}$$

which corresponds to the air density.

Contour branches with parallel unidirectional currents have to twist, so ball lightning contours are gradually opening losing the charge. Therefore, ball lightning has a sort of

an electrostatic tail behind. The maximum lifetime of ball lightning can be determined in a similar way as the neutron lifetime [2], i.e., as the *time constant* of the contour deformation (the ratio of the contour's characteristic dimension to the circumferential velocity):

$$\tau_{lim} = \frac{l_k}{V_{0i}} = 327 \text{ sec.} \tag{20}$$

When crossing the initial surface of our world, an open contour (vortical tube) actually forms an elementary charge (according to Wheeler). It can be assumed that the physical basis of ball lightning is formed by electrons. Their fermionic part is arranged into corresponding structures observable in the form of a fireball, while their bosonic parts converge in the centre of the ball going to the additional dimension (Y area) [2].

Let us also determine the capacity, electrostatic energy, and size required for ball lightning with q_{lim} charge. Paper [**?**] found a connection between new electrical units in a *Coulomb-free* form and SI-system units. It was shown that the mass of electrons 2.90×10^{-6} kg on the capacitor plates corresponded to one Farad. Velocity of 587 m/sec corresponded to 1 Volt, with the electrostatic capacity of the surface at which the charge begins to flow spontaneously into the external environment being $U_m = 511,000$ V.

Thus, ball lightning has the following capacity:

$$C = \frac{zm_e}{2.90\times10^{-6}} = 4.89\times10^{-6} \text{ F}, \tag{21}$$

with the same result in the SI system:

$$C = \frac{q_{lim}}{511,000} = 4.89\times10^{-6} \text{ F}, \tag{22}$$

and the maximum electrostatic energy of ball lightning being

$$E_m = \frac{1}{2} CU_m^2 = 6.39\times10^5 \text{ J}. \tag{23}$$

To have such a capacity, ball lightning must have a multilayer structure, e.g., the structure of a multilayer spherical capacitor. Paper [5] shows that the average distance between unit charges of a charged sphere with R radius is $\pi(Rr_e)^{1/2}$. Let us assume that the average linear dimension between the charges in the volume of ball lightning is the same. Then we can determine the size of ball lightning through the following equation:

$$\left(\frac{\frac{4}{3}\pi R^3}{z}\right)^{1/3} = \pi\left(Rr_e\right)^{1/2}, \tag{24}$$

therefore

$$R = \left(\frac{3\pi^2 z}{4}\right)^{2/3} r_e = 0.067 \text{ m}. \tag{25}$$

Let us determine the temperatures of the nucleus and the outer shell with the assumption that the radiation of ball lightning is the radiation of a blackbody. If the total energy is evenly lost over the lifetime of ball lightning, the average radiation power shall be as follows:

$$N = \frac{E_m}{\tau_{lim}} = 1950 \text{ W}, \tag{26}$$

then

$$T = \left(\frac{N}{\sigma S}\right)^{1/4}, \tag{27}$$

where σ is a Stefan-Boltzmann constant equal to 5.67×10^{-8} Wm^{-2} $(^\circ\text{K})^{-4}$; and S is the area of the spherical surface of ball lightning. Taking l_{min} as the nucleus diameter and l_{max} as the diameter of the outer shell, we calculate the respective areas S and determine their temperatures using formula (27) — 8.500°K and 724°K. External appearance of ball lightning, its behaviour, and results of its effect on the environment are extremely varied. Given its unpredictability, it is rarely possible to obtain objective instrumental data on ball lightning.

In his paper [6], Mikhail Dmitriev — a chemist having vast experience in working with low-temperature plasma — describes an encounter with ball lightning and an attempt to make a chemical analysis of ionized air behind it. Based on the analysis results, the author estimated the potential of ball lightning discharge at 300–400 kV. The temperature, degree of ionization, and concentration of charged particles in ball lightning was estimated at 1.14×10^{17} per cm^3 judging on its glow. It is easy to calculate that, in accordance with the proposed model and given such a concentration, the estimated diameter of ball lightning with z unit charges shall be 6.4 cm, which corresponds to its typical size. This means that the discharge potential and charge concentration of real ball lightning encountered by Dmitriev are consistent with the estimated model.

Since ball lightning does not consist of atoms and molecules, it does not interact with molecules of other media. This explains its ability to penetrate through obstacles and move against the wind, but actively respond to electric and magnetic fields at the same time.

Finally, it should be noted that people often associate ball lightning with a living being. Let us assume that life can be organized on another material basis. Then, indeed, given the number of unit elements ($z = 1.56\times10^{19}$) and complexity of their packing in the volume of ball lightning, it is appropriate to draw an analogy with a DNA strand, which is two meters long, packed in a microscopic cell nucleus, and contains information about the structure and behaviour of a living organism.

4 Conclusion

Thus, model ball lightning is a ball with its size ranging from 0.14 to 20 cm (its typical diameter is 13.4 cm), having density of no more than 1.17 kg per m^3, glow temperature of 724 to 8,500°K, and energy of 639 kJ concentrated in a small volume in the form of an electrostatic charge with 511 kV potential.

During the lifetime of ball lightning (up to 6.5 minutes), it is constantly losing the charge leaving an ionised trail behind. Ball lightning is able to penetrate obstacles.

In general, eyewitness accounts are in good agreement with the calculated characteristic parameters of the model object. Of course, some phenomena of real ball lightning fall outside the scope of the obtained characteristic parameters. At least this is due to the fact that its charge can be formed by not only electrons, but also by ions, and the evolutionary parameter ε may exceed the unit.

The ball lightning phenomenon and its complete internal organization can only be understood on the basis of an appropriate theory. However, from a phenomenological point of view, this model of ball lightning is in good agreement with the real object by its appearance and its basic aspects, and can serve as the basis for such a theory.

Submitted on April 28, 2016 / Accepted on May 20, 2016

References

1. Belyakov, A.V.: Charge of the electron, and the constants of radiation according to J.A.Wheeler's geometrodynamic model. *Progress in Physics*, 2010, vol. 6, issue 4, 90–94.
2. Belyakov A.V. Macro-analogies and gravitation in the micro-world: further elaboration of Wheeler's model of geometrodynamics. *Progress in Physics*, 2012, vol. 8, issue 2, 47–57.
3. Belyakov A.V. Evolution of stellar objects according to J. Wheeler's geometrodynamic concept. *Progress in Physics*, 2013, vol. 9, issue 1, 25–40.
4. Ignatov B.A. Ball lightning as a quasi-particle's child. In the book series: *Club of Fundamental Natural-Science Ideas. Hypotheses. Judgments. Opinions. Speculations*. Issue 1. Fenid, 1990.
5. Belyakov A.V. On the uniform dimension system. Is there the necessity for Coulomb? *Progress in Physics*, 2013, vol. 9, issue 3, 142–143.
6. Dmitriev M.T. Ball lightning nature. *Priroda*, June 1967, 98–106 (*in Russian*).

LETTERS TO PROGRESS IN PHYSICS

Dialogue Concerning the Two Chief World Views

Craig Alan Feinstein
2712 Willow Glen Drive, Baltimore, MD 21209, USA. E-mail: cafeinst@msn.com

In 1632, Galileo Galilei wrote a book called *Dialogue Concerning the Two Chief World Systems* which compared the new Copernican model of the universe with the old Ptolemaic model. His book took the form of a dialogue between three philosophers, Salviati, a proponent of the Copernican model, Simplicio, a proponent of the Ptolemaic model, and Sagredo, who was initially open-minded and neutral. In this paper, I am going to use Galileo's idea to present a dialogue between three modern philosophers, Mr. Spock, a proponent of the view that $\mathsf{P} \neq \mathsf{NP}$, Professor Simpson, a proponent of the view that $\mathsf{P} = \mathsf{NP}$, and Judge Wapner, who is initially open-minded and neutral.

Since 2006, I have published four proofs that $\mathsf{P} \neq \mathsf{NP}$ [5–8]. Yet at the present time, if one asks the average mathematician or computer scientist the status of the famous P versus NP problem, he or she will say that it is still open. In my opinion, the main reason for this is because most people, whether they realize it or not, believe in their hearts that $\mathsf{P} = \mathsf{NP}$, since this statement essentially means that problems which are easy to state and have solutions which are easy to verify must also be easy to solve. For instance, as a professional magician, I have observed that most laymen who are baffled by an illusion are usually convinced that the secret to the illusion either involves extraordinary dexterity or high technology, when in fact magicians are usually no more dexterous than the average layman and the secrets to illusions are almost always very simple and low-tech; as the famous designer of illusions, Jim Steinmeyer, said, "Magicians guard an empty safe" [13]. The thinking that extraordinary dexterity or high technology is involved in a magician's secret is, in my opinion, due to a subconscious belief that $\mathsf{P} = \mathsf{NP}$, that problems which are difficult to solve and easy to state, in this case "how did the magician do it?", must have complex solutions.

I have had many conversations in which I have tried to convince all types of people, from Usenet trolls to graduate students to professors to famous world-class mathematicians, that $\mathsf{P} \neq \mathsf{NP}$ with very little success; however, I predict that there will soon come a day when the mainstream mathematics and computer science community will consider people who believe that $\mathsf{P} = \mathsf{NP}$ to be in the same league as those who believe it is possible to trisect an angle with only a straightedge and compass (which has been proven to be impossible) [14].

I got the idea to write this paper after I learned of Galileo's book *Dialogue Concerning the Two Chief World Systems* [4], which presents a dialogue between three philosophers, Salviati, a proponent of the new Copernican model, Simplicio, a proponent of the old Ptolemaic model, and Sagredo, who was initially open-minded and neutral. The dialogue that follows is a dialogue between three modern philosophers, Mr. Spock, a proponent of the view that $\mathsf{P} \neq \mathsf{NP}$, Professor Simpson, a proponent of the view that $\mathsf{P} = \mathsf{NP}$, and Judge Wapner, who is initially open-minded and neutral. Professor Simpson, who is a fictitious anglicized straw man character like Simplicio, is a composite of many of the people whom I have had discussions with over the years about the P versus NP problem. He presents many challenges and questions, all of which have been raised before by real people, that Mr. Spock, the epitome of truth and logic, attempts to answer. And Judge Wapner, the epitome of open-mindedness and fairness, always listens to both sides of their arguments before drawing conclusions.

Spock: Yesterday we discussed the P versus NP problem [2, 3] and agreed that it is a problem of not only great philosophical importance, but also it has practical implications. We decided to look at a proof that $\mathsf{P} \neq \mathsf{NP}$ offered by Craig Alan Feinstein in a letter entitled "A more elegant argument that $\mathsf{P} \neq \mathsf{NP}$" [8]. The proof is surprisingly short and simple:

Proof: Consider the following problem: Let $\{s_1, \ldots, s_n\}$ be a set of n integers and t be another integer. Suppose we want to determine whether there exists a subset of $\{s_1, \ldots, s_n\}$ such that the sum of its elements equals t, where the sum of the elements of the empty set is considered to be zero. This famous problem is known as the SUBSET-SUM problem.

Let $k \in \{1, \ldots, n\}$. Then the SUBSET-SUM problem is equivalent to the problem of determining whether there exist sets $I^+ \subseteq \{1, \ldots, k\}$ and $I^- \subseteq \{k+1, \ldots, n\}$ such that

$$\sum_{i \in I^+} s_i = t - \sum_{i \in I^-} s_i.$$

There is nothing that can be done to make this equation simpler. Then since there are 2^k possible expressions on the left-hand side of this equation and 2^{n-k} possible expressions on the right-hand side of this equation, we can find a lower-bound for the worst-case running-time of an algorithm that solves the SUBSET-SUM problem by minimizing $2^k + 2^{n-k}$ subject to $k \in \{1, \ldots, n\}$.

When we do this, we find that $2^k+2^{n-k} = 2^{\lfloor n/2 \rfloor}+2^{n-\lfloor n/2 \rfloor} = \Theta(\sqrt{2^n})$ is the solution, so it is impossible to solve the SUBSET-SUM problem in $o(\sqrt{2^n})$ time; thus, because the Meet-in-the-Middle algorithm [10,11,15] achieves a running-time of $\Theta(\sqrt{2^n})$, we can conclude that $\Theta(\sqrt{2^n})$ is a tight lower-bound for the worst-case running-time of any deterministic and exact algorithm which solves SUBSET-SUM. And this conclusion implies that $\mathsf{P} \neq \mathsf{NP}$. □

To me, Feinstein's proof is not only logical but elegant too. Also, his conclusion is confirmed by history; just as Feinstein's theorem retrodicts, no deterministic and exact algorithm that solves SUBSET-SUM has ever been found to run faster than the Meet-in-the-Middle algorithm, which was discovered in 1974 [10, 15].

Simpson: But there is an obvious flaw in Feinstein's "proof": Feinstein's "proof" only considers a very specialized type of algorithm that works in the same way as the Meet-in-the-Middle algorithm, except that instead of sorting two sets of size $\Theta(\sqrt{2^n})$, it sorts one 2^k-size set and one 2^{n-k}-size set. Under these restrictions, I would agree that the Meet-in-the-Middle algorithm is the fastest deterministic and exact algorithm that solves SUBSET-SUM, but there are still many possible algorithms which could solve the SUBSET-SUM problem that the "proof" does not even consider.

Wapner: Professor Simpson, where in Feinstein's proof does he say that he is restricting the algorithms to the class of algorithms that you mention?

Simpson: He does not say so explicitly, but it is obviously implied, since there could be algorithms that get around his assertion that the minimum number of possible expressions on both sides is $\Theta(\sqrt{2^n})$.

Spock: How do you know that there could be such algorithms?

Simpson: I do not know, but the burden of proof is not on me; it is on Feinstein. And he never considers these types of algorithms.

Wapner: It is true that Feinstein never explicitly considers algorithms which work differently than the Meet-in-the-Middle algorithm, and the burden of proof is on Feinstein to show that these types of algorithms cannot run any faster than $\Theta(\sqrt{2^n})$ time.

Spock: Professor Simpson, is the burden of proof on Feinstein to consider in his proof algorithms which work by magic?

Simpson: No, only algorithms that are realistic.

Spock: Then why would you think that algorithms that get around the assertion that the minimum total number of possible expressions on both sides is $\Theta(\sqrt{2^n})$ are realistic?

Simpson: I do not know, but the burden of proof is not on me; it is on Feinstein.

Spock: Have you considered the fact that an algorithm which determines in $o(\sqrt{2^n})$-time whether two sets of size $\Theta(\sqrt{2^n})$ have a nonempty intersection *must* work by magic, unless there is a way to mathematically reduce the two sets into something simpler?

Wapner: Yes, I see your point; the minimum total number of possible expressions on each side of the SUBSET-SUM equation puts a natural restriction on the time that an algorithm must take to solve the SUBSET-SUM problem.

Simpson: But how do you know it is impossible to reduce the SUBSET-SUM problem into something simpler, so that the number of possible expressions on both sides is $o(\sqrt{2^n})$?

Spock: Simple algebra. Try to simplify the SUBSET-SUM equation above. You cannot do it. The best you can do is manipulate the equation to get $\Theta(\sqrt{2^n})$ expressions on each side.

Simpson: I'll agree that you cannot do it algebraically, but what about reducing the SUBSET-SUM problem to the 3-SAT problem in polynomial-time? This can be done since 3-SAT is NP-complete. If there is an algorithm that can solve 3-SAT in polynomial-time, then it would also be able to solve SUBSET-SUM in polynomial-time, contradicting Feinstein's lower-bound claim of $\Theta(\sqrt{2^n})$.

Spock: But this is magical thinking. If a problem is shown to be impossible to solve in polynomial time, then reducing the problem to another problem in polynomial-time will not change the fact that it is impossible to solve the first problem in polynomial time; it will only imply that the second problem cannot be solved in polynomial time.

Wapner: Spock is right about this. Do you have any other objections to Feinstein's argument?

Simpson: I have many objections. For instance, Feinstein's argument can be applied when the magnitudes of the integers in the set $\{s_1, \ldots, s_n\}$ and also t are assumed to be bounded by a polynomial to "prove" that it is impossible to solve this modified problem in polynomial-time. But it is well-known that one can solve this modified problem in polynomial-time.

Spock: But Feinstein's argument in fact *cannot* be applied in such a circumstance, because there would only be a polynomial number of possible values on each side of the equation,

even though the total number of possible expressions on each side is exponential. Feinstein's argument implicitly uses the fact that the total number of possible values on each side of the SUBSET-SUM equation is usually of the same order as the total number of possible expressions on each side, when there is no restriction on the magnitude of the integers in the set $\{s_1, \ldots, s_n\}$ and also t.

Simpson: Then here is a better objection: Suppose the set $\{s_1, \ldots, s_n\}$ and also t consist of vectors in $\mathbb{Z}_2^m$ for some positive integer m, instead of integers. Then one could use the same argument that Feinstein uses to "prove" that it is impossible to determine in polynomial-time whether this modified SUBSET-SUM equation has a solution, when in fact one can use Gaussian elimination to determine this information in polynomial-time.

Spock: Feinstein's argument would not apply to this situation precisely because one can reduce the equation

$$\sum_{i \in I^+} s_i = t - \sum_{i \in I^-} s_i.$$

to a simpler set of equations through Gaussian elimination. But when the set $\{s_1, \ldots, s_n\}$ and also t consist of integers, nothing can be done to make the above equation simpler, so Feinstein's argument is applicable.

Simpson: OK, then how would you answer this? Consider the Diophantine equation:

$$s_1 x_1 + \ldots + s_n x_n = t,$$

where x_i is an unknown integer, for $i = 1, \ldots, n$. One could use the same argument that Feinstein uses to "prove" that it is impossible to determine in polynomial-time whether this equation has a solution, when in fact one can use the Euclidean algorithm to determine this information in polynomial-time.

Spock: But again Feinstein's argument would not apply to this Diophantine equation, precisely because this Diophantine equation can be reduced via the Euclidean algorithm to the equation,

$$\gcd(s_1, \ldots, s_n) \cdot z = t,$$

where z is an unknown integer. And it is easy to determine in polynomial-time whether this equation has an integer solution by simply testing whether t is divisible by $\gcd(s_1, \ldots, s_n)$. No such reduction is possible with the SUBSET-SUM equation.

Simpson: The Euclidean algorithm is a clever trick that has been known since ancient times. But how do I know that another clever trick cannot be found to reduce the SUBSET-SUM equation to something simpler? Like for instance, if I take the greatest common denominator of any subset of $\{s_1, \ldots, s_n\}$ and it does not divide t, then I can automatically rule out many solutions to SUBSET-SUM, all at once.

Spock: But such a clever trick does not always work; what if the gcd *does* divide t? The P versus NP problem is a problem about the *worst-case* running-time of an algorithm, not whether there are clever tricks that can be used to speed up the running-time of an algorithm in some instances. Feinstein's proof only considers the *worst-case* running-time of algorithms which solve SUBSET-SUM.

Wapner: Also, it is simple high school algebra that it is impossible to make the SUBSET-SUM equation simpler than it is: Whenever one decreases the number of possible expressions on one side of the equation, the number of possible expressions on the other side increases. Mathematicians can be clever, but they cannot be clever enough to get around this fact.

Simpson: OK, but what about the fact that Feinstein never mentions in his proof the model of computation that he is considering? To be an valid proof, this has to be mentioned.

Spock: Feinstein's proof is valid in any model of computation that is realistic enough so that the computer cannot solve an equation with an exponential number of possible expressions in polynomial-time, unless it is possible to reduce the equation to something simpler.

Simpson: Or what about the fact that Feinstein never mentions in his paper the important results that one cannot prove that $\mathsf{P} \neq \mathsf{NP}$ through an argument that relativizes [1] or through a natural proof [12]?

Spock: Feinstein's proof does not relativize, because it implicitly assumes that the algorithms that it considers do not have access to oracles, and Feinstein's proof is not a natural proof, since it never even deals with the circuit complexity of boolean functions.

Simpson: What about the 2010 breakthrough by Howgrave-Graham and Joux [9] which gives a probabilistic algorithm that solves SUBSET-SUM in $o(\sqrt{2^n})$ time? I realize that the P versus NP problem is not about probabilistic algorithms, but what if their algorithm can be derandomized and solved in $o(\sqrt{2^n})$ time?

Spock: The algorithm by Howgrave-Graham and Joux does not in fact solve SUBSET-SUM, because it cannot determine for certain when there is no solution to a given instance of SUBSET-SUM; it can only output a solution to SUBSET-SUM in $o(\sqrt{2^n})$ time with high probability when a solution

exists. Furthermore, even if their algorithm can be derandomized, this does not guarantee that it will run in $o(\sqrt{2^n})$ time. And Feinstein has already proven that such a deterministic and exact algorithm is impossible.

Wapner: Are there any more objections to Feinstein's argument?

Simpson: I have no more specific objections. But the fact that the P versus NP problem has been universally acknowledged as a problem that is very difficult to solve and Feinstein's "proof" is so short and simple makes it almost certain that it is flawed. The fact that I could not give satisfactory responses to Spock's arguments does not mean that Feinstein is correct; Feinstein's proof has been out on the internet for a few years now, and still the math and computer science community as a whole does not accept it as valid. Hence, I believe that they are right and that Feinstein is wrong.

Wapner: Professor Simpson, isn't your reason for not believing Feinstein's proof the same reason Feinstein suggested for why most people do not believe his proof? Because most people believe in their hearts that P=NP, that problems which are difficult to solve and easy to state, in this case the P versus NP problem, cannot have short and simple solutions?

Spock: Indeed it is.

Wapner: And yes indeed, I am convinced that Feinstein's proof is valid and that $\mathsf{P} \neq \mathsf{NP}$.

Submitted on May 22, 2016 / Accepted on May 24, 2016

References

1. Baker T.P., Gill J., and Solovay R. Relativizations of the P ≠ NP Question. *SIAM Journal on Computing*, 1975, v. 4(4), 431–442.
2. Bovet P.B. and Crescenzi P. Introduction to the Theory of Complexity. Prentice Hall, 1994.
3. Cormen T.H., Leiserson C.E., and Rivest R.L. Introduction to Algorithms. McGraw-Hill, 1990.
4. Galilei G. Dialogue Concerning the Two Chief World Systems. University of California Press, 1953.
5. Feinstein C.A. Complexity science for simpletons. *Progress in Physics*, 2006, v. 2, issue 3, 35–42.
6. Feinstein C.A. An elegant argument that $P \neq NP$. *Progress in Physics*, 2011, v. 7, issue 2, 30–31.
7. 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.
8. Feinstein C.A. A more elegant argument that $P \neq NP$. *Progress in Physics*, 2012, v. 8, issue 1, L10.
9. Howgrave-Graham N., and Joux A. New generic algorithms for hard knapsacks. *IACR Cryptology ePrint Archive*, 2010, 189.
10. Horowitz E. and Sahni S. Computing partitions with applications to the Knapsack Problem. *Journal of the ACM*, 1974, v. 2l, no. 2, 277–292.
11. Menezes A., van Oorschot P., and Vanstone S. Handbook of Applied Cryptography. CRC Press, 1996.
12. Razborov A. and Rudich S. Natural proofs. *Journal of Computer and System Sciences*, 1997, v. 55, issue 1, 24–35.
13. Steinmeyer J. Hiding the Elephant: How Magicians Invented the Impossible and Learned to Disappear. Da Capo Press, 2004.
14. Angle Trisection. From MathWorld — A Wolfram Web Resource. http://mathworld.wolfram.com/AngleTrisection.html
15. Woeginger G.J. Exact algorithms for NP-hard problems. In: *Lecture Notes in Computer Science*, Springer-Verlag Heidelberg, 2003, v. 2570, 185–207.

Criteria for Aerial Locomotion in Exoplanetary Atmospheres: Revisiting the Habitable Zone for Flying Lifeforms

Robin James Spivey

Bangor University, Deiniol Road, Bangor, Gwynedd, Great Britain
E-mail: y.gofod@gmail.com

Liquid water is widely regarded as a hallmark of planetary habitability but, whilst its presence may be a prerequisite for life, aerial locomotion imposes additional constraints on the somewhat over-simplistic concept of a circumstellar habitable zone. Could animals of comparable physiology to birds be envisaged sustaining flight without environmental assistance on super-Earth planets of terrestrial density? A quantitative evaluation of flight athleticism in avian species provides the basis for extrapolation here. At constant atmospheric fraction, assuming a plentiful supply of combustible gas, the "aerial locomotion zone" would be restricted to planets $\lesssim$6.86 $M_\oplus$. However, due to the inevitable thermal impediments at higher altitudes, it is conceivable that the majority of the Earth's avian species could evolve sufficient athleticism for flight on temperate isoatmospheric planets of up to 15 $M_\oplus$, without adjustments in body mass.

1 Introduction

Birds, bats, insects and pterosaurs have independently surmounted the challenges of actively-powered flight [1], perhaps during hyperoxic episodes in the Earth's history [2]. Avian species span some four orders of magnitude in body mass [3] yet birds of all sizes undertake arduous seasonal migrations [4,5]. Flight is a complex and intrinsically dangerous activity especially in arboreal environments, over mountainous terrain, in regions where birds of prey are prevalent or during unfavourable weather [6]. Thus, there is a need for sophisticated neural control [7]. Exoplanet discoveries continue apace [8,9] and NASA's Kepler mission has already established that those of 1–2 Earth radii ("super-Earths") are remarkably abundant [10]. Neutrally buoyant aquatic animals are immune to changes in gravity and land animals can evolve sturdier bones or additional legs to cope with conditions on more massive planets. However, the feasibility of environmentally unassisted flight in stronger gravitational fields is clearly an intricate issue meriting more detailed scrutiny.

The Earth's oxygenated air provides birds not only with a breathable atmosphere but also a medium for generating propulsion and weight support during flight [11]. Consequently, gravity, atmospheric density and the chemical composition of an atmosphere influence the prospects for aerial locomotion. There is no evidence that the laws of physics vary either with time or location, so animals that are as anatomically and physiologically well-adapted to flight as any living here could have evolved elsewhere in the universe. This analysis therefore commences by evaluating the athleticism of Earth's avian species during environmentally unassisted horizontal flight. The limits of flight athleticism on Earth are then used as a basis for extrapolation to different planetary environments, leading to criteria that are likely to be satisfied if circumplanetary atmospheres are compatible with flight.

2 Flight power and athleticism

The following analysis concerns flying animals capable of supporting their own weight in still air conditions, building upon an established result from aerodynamic theory pertaining to hovering flight [12]. If a bird's wings have combined area A_{wing} and the air they sweep is on average accelerated to a downward velocity v_a then the volume of air being swept in unit time is $v_a A_{wing}$. In an atmosphere of density ρ, the mass of this parcel of air is $m_a = \rho v_a A_{wing}$ and so the rate of change of momentum in the air is $m_a v_a = \rho v_a^2 A_{wing}$. For a bird of body mass m_b, Newton's second law requires that this equals the bird's weight $m_b g$ which allows the downward velocity of the air to be obtained as $v_a = \sqrt{m_b g/\rho A_{wing}}$. The power required during hovering is the rate at which kinetic energy is imparted to the air

$$P_{hov} = \frac{m_a v_a^2}{2} = \frac{\rho A_{wing} v_a^3}{2} = \frac{1}{2}\sqrt{\frac{m_b^3 g^3}{\rho A_{wing}}}. \qquad (1)$$

Providing only a small fraction of the power relating to forward horizontal flight, P_f, is required to overcome the drag associated with forward motion, it can be argued that P_f and P_{hov} should scale almost identically. If, furthermore, avian anatomy scales isometrically then $A_{wing} \propto m_b^{2/3}$ and

$$P_f \propto \left(\frac{m_b^7 g^9}{\rho^3}\right)^{1/6}. \qquad (2)$$

For an individual animal this simplifies to $P_f \propto \sqrt{g^3/\rho}$, a term which concisely encapsulates environmental conditions. Thus, flight becomes more challenging on planets with stronger gravitational fields and reduced atmospheric densities [13]. On Earth, flying birds and flightless birds are delineated by the boundary $\sqrt{g^3/\rho}$ =27.7 $m^3\,s^{-3}\,kg^{-1/2}$. Departures from isometry are likely [14] and the allometrically

neutral relationship $A_{wing} \propto m_b^{2/3}$ is only marginally compatible with empirical data – actual wing measurements suggest $A_{wing} \propto m_b^{0.780\pm0.112}$ [15] . This implies the following modification involving an exponent $\alpha = 0.110 \pm 0.056$,

$$P_f \propto m_b^{1+\alpha} \sqrt{\frac{g^3}{\rho}}. \tag{3}$$

Since $\alpha > 0$, mass-specific flight power, P_f/m_b, generally increases with body mass [12]. A quantity χ is now introduced which is directly proportional to the mass-specific flight power needed to fly horizontally in still air. It is adopted as a proxy for flight athleticism and defined as

$$\chi = \left(\frac{m_b}{\tilde{m}}\right)^{\alpha} \sqrt{\frac{g^3}{\rho}} \tag{4}$$

where $\tilde{m}$ is a fiducial mass term used for normalisation and can be arbitrarily chosen. In particular, heavier animals capable of flight in hypodense air would score well on this measure. At a similar airspeed, aerodynamic drag is of less concern to large birds, in keeping with the earlier assumption that $P_f \propto P_{hov}$.

3 Aeronautical limits

Avian lungs utilise a cross-current airflow assisted by a complementary vasculature allowing for efficient gas exchange [16], advantageous during high altitude flight where the partial pressure of oxygen is reduced. At least three species appear to be capable of entirely self-powered flight 7000 m or more above sea level. An iconic example is the bar-headed goose, *Anser indicus*, whose seasonal migrations involve navigating the Himalaya [17] and the prominent obstacle of the Tibetan plateau. Having been satellite-tracked at 7290 m [18] they are more tolerant of hypoxia than brent geese, *Branta bernicla*, which have difficulty crossing the Greenland icecaps at altitudes of 2500 m [19]. Despite a body mass placing them in the 98$^{\text{th}}$ percentile of bird species [3], they have also been observed flying in formation at almost 8000 m by mountaineers climbing the Annapurna massif [20]. A number of cardiovascular, pulmonary, morphological and biochemical adaptation mechanisms could be responsible for this striking athleticism including high ventilation rates [21], relative immunity to respiratory alkalosis and haemoglobin of superior O_2 affinity, higher cardiac output [17] and tissue enhancements such as cardiac hypertrophy, greater capillary density and mitochondrial abundance [22].

Alpine choughs, *Pyrrhocorax graculus* also inhabit the Himalaya. Nesting as high as 6500 m [17], they have been known to follow climbers on Everest at altitudes approaching 8200 m – within the mountaineering "death zone". Small birds such as choughs readily take to the air but swans are much larger and typically require 15–20 wingbeats to become aloft when taking off from water, even though they can obtain some acceleration and weight support from webbed feet. On becoming airborne they continue to gain speed and gradually start to ascend, necessitating continued effort [23]. Thus, unlike smaller birds for which a short period of anaerobic exertion is adequate for take-off, swans must demonstrate aerobic athleticism at the commencement of each flight. This applies also to juveniles – cygnets only start to fly at 4–5 months of age. The athleticism demanded by take-off may confer upon swans an ability to sustain high altitude flight, even if they are not ecologically coerced to do so. Lowland species may be incapable of take-off in hypodense air but that does not preclude, *per se*, an ability to fly high – even though swans tend not to during migration [24]. In still air conditions, flying low in dense air facilitates flight – in accordance with (4). However, strong tailwinds capable of drastically curtailing migration times and total energy expenditure are sometimes available, especially at higher altitudes. During lengthier migratory flights, the additional costs of ascent and high altitude cruising can easily be fully recovered. In the cold and featureless seascape of the north Atlantic, which is neither conducive to the generation of strong thermals nor orographic updrafts, a flock of some 30 whooper swans, *Cygnus cygnus*, was detected in 1967 by radar then visually identified by a pilot to be flying at 8200 m with a ground-speed of 38 m s^{-1} towards the end of a ∼1000 km migration from Iceland to the UK [25].

The air density at 8200 m is 0.513 kg m^{-3}. Setting $\tilde{m} = m_w = 11$ kg, the mass of a whooper swan, and making allowances for variations in α, the maximum value of χ at which flight is possible at this altitude is 42.8 m^3 s^{-3} kg$^{-1/2}$ for whooper swans, 34.1–39.8 m^3 s^{-3} kg$^{-1/2}$ for bar-headed geese and 22.4–34.6 m^3 s^{-3} kg$^{-1/2}$ for Alpine choughs. Results for various species are presented in Fig. 1. Whooper swans appear to top the list for avian athleticism making them well-suited for astrobiological extrapolations. To compete, bar-headed geese, would need to be capable of flight at altitudes of 9.4–11.7 km, which seems unrealistically high [18].

4 Planetary environments

The radii, R, of terrestrial super-Earths are expected to scale with M, the planet mass, as $R \propto M^{0.274}$ [26]. Hence, surface gravity, g_s, should scale as $g_s \propto M^{0.452}$. The effective increase in sea level on a super-Earth planet with a similar water content to Earth can be estimated from the relationship plotted in Fig. 2. It is also relevant to mention that enhanced gravity tends to attenuate topographical features such as mountains and ridges. Super-Earth planets are variously taken to have a mass of 1–10 $M_\oplus$ or a radius of 1–2 $R_\oplus$ (1–12 $M_\oplus$) where the subscript $\oplus$ denotes the Earth. This analysis considers the slightly expanded range 1–15 $M_\oplus$ in order to encompass the largest planets capable of possessing hexagonally close-packed iron at their core [27, 28]. On Earth, a whooper swan can fly in air of density as low as $\rho_w = 0.513$ kg m^{-3}. Since athleticism is not an environmental variable, the min-

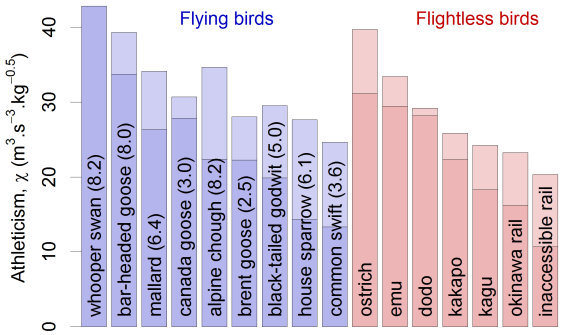

Fig. 1: Flight athleticism, χ, for various species. Estimated maximum altitudes are given in km (in parentheses) for unassisted flight. Selected results are also provided for flightless birds assuming (mostly with undue optimism) that they might be capable of flying in air slightly denser than that of sea level. Lightly shaded areas represent the uncertainty in the allometric scaling exponent, $0.054 < \alpha < 0.166$, using a fiducial mass $\tilde{m}$=11 kg.

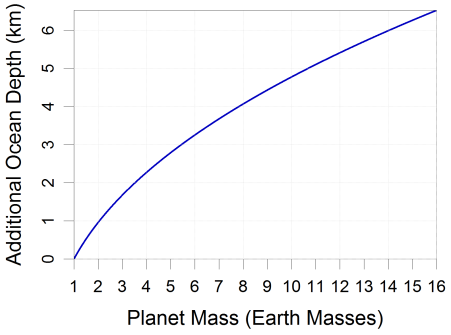

Fig. 2: Planets more massive than Earth but with an identical water fraction ($V_{H_2O} \propto M$) would have somewhat deeper oceans, the additional depth (in km) being at least 2.6 $[(M/M_\oplus)^{0.452} - 1]$ depending on topography. However, if planetary water is exclusively delivered from space via comets and asteroids whose spatial distribution varies little with galactic location, one would anticipate ocean depths to be largely independent of planet mass.

imum air density required by whooper swans on other planets is $\rho_{min} = \rho_w (M/M_\oplus)^{1.356}$. For a 15 Earth-mass planet, ρ_{min}=20.18 kg m^{-3}.

As with discussions of circumstellar habitable zones, some simplifying assumptions are helpful. Due to uncertainties such as cloud cover, humidity levels and fluctuations in atmospheric heating due to planetary rotation, an isothermal model atmosphere is adopted. In hydrostatic equilibrium the ideal gas law predicts that air density, ρ, is proportional to air pressure, $p = \beta\rho$. According to the International Standard Atmosphere, β has a value of about 82714 m^2 s^{-2} for a temperature of 15°C at sea level where ρ =1.225 kg m^{-3}. Gravity is taken to be insensitive to changes in altitude, z. By considering the weight of a thin horizontal layer of air,

$$\frac{dp}{dz} = \frac{dp}{d\rho} \times \frac{d\rho}{dz} = \beta \frac{d\rho}{dz} = -\rho g_s. \tag{5}$$

The air density at height h is obtained by integrating from $z = 0$ to $z = h$,

$$\beta (\ln \rho - \ln \rho_s) = -g_s h, \tag{6}$$

where ρ_s represents the air density at the surface. Thus, $\rho = \rho_s \exp(-g_s h/\beta)$ and the total mass contained by the atmosphere below height h is

$$M_h = 4\pi R^2 \int_0^h \rho(z)\, dz = \frac{4\pi\beta\rho_s R^2}{g_s} \left[1 - \exp\left(\frac{-g_s h}{\beta}\right)\right]. \tag{7}$$

M_h converges as $h \to \infty$ to yield the total mass of the entire atmosphere,

$$M_{atm} = \frac{4\pi\beta\rho_s R^2}{g_s}. \tag{8}$$

For an isothermal atmosphere, under the assumption of spherical symmetry, half the air mass lies below a scale height $\hat{h}$ given by

$$\hat{h} = \frac{\beta \ln 2}{g_s} = \frac{\beta \ln 2}{g_\oplus} \left(\frac{M_\oplus}{M}\right)^{0.452}. \tag{9}$$

This expression is entirely independent of ρ_s. Plots of surface gravity, planetary radius and atmospheric scale height against planetary mass are provided in Fig. 3.

5 Criteria for aerial locomotion

From (8) we have $\rho_s = g_s M_{atm}/4\pi\beta R^2$. Recalling that $g_s \propto M^{0.452}$ and $R \propto M^{0.274}$,

$$\rho_s = \frac{g_\oplus M_{atm} (M/M_\oplus)^{0.452}}{4\pi\beta R_\oplus^2 (M/M_\oplus)^{0.548}} \propto M^{0.904} \left(\frac{M_{atm}}{M}\right). \tag{10}$$

Since $g_s = g_\oplus (M/M_\oplus)^{0.452}$, the quantity $\sqrt{g_s^3/\rho_s}$, a factor previously found to be proportional to the power required by flight, can be expressed as follows

$$\begin{aligned} \frac{g_s^3}{\rho_s} &= \frac{4\pi\beta R_\oplus^2 (M/M_\oplus)^{0.096}}{g_\oplus M_{atm}} \left[g_\oplus (M/M_\oplus)^{0.452}\right]^3 \\ &= \frac{4\pi\beta R_\oplus^2 g_\oplus^2}{M_{atm}} \left(\frac{M}{M_\oplus}\right)^{1.452}, \end{aligned} \tag{11}$$

$$\sqrt{\frac{g_s^3}{\rho_s}} = \gamma \left(\frac{M}{M_\oplus}\right)^{0.226} \sqrt{\frac{M}{M_{atm}}} \tag{12}$$

where $\gamma = 2 g_\oplus R_\oplus \sqrt{\pi\beta/M_\oplus} = 0.026$ m^3 s^{-3} kg$^{-1/2}$. The maximum mass of an isoatmospheric planet (i.e. having a ratio M_{atm}/M identical to Earth's) that is compatible with flight for

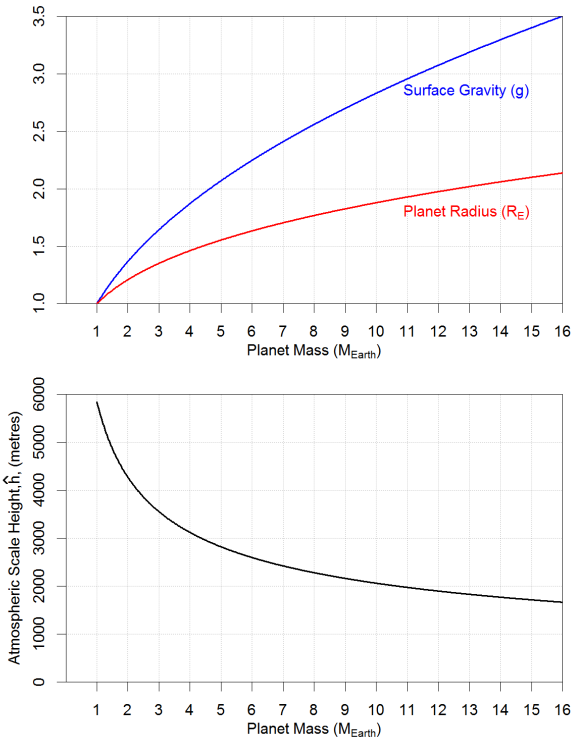

Fig. 3: Upper panel: planetary radius and surface gravity obey simple power law relationships according to planetary mass, $R \propto M^{0.274}$ and $g_s \propto M^{0.452}$ respectively. Lower panel: the scale height of the atmosphere, $\hat{h} = \beta \ln 2/g_s$, is independent of the surface air density and hence total mass of the atmosphere. It decreases for larger planets since a higher surface gravity is better able to confine the atmosphere close to the surface.

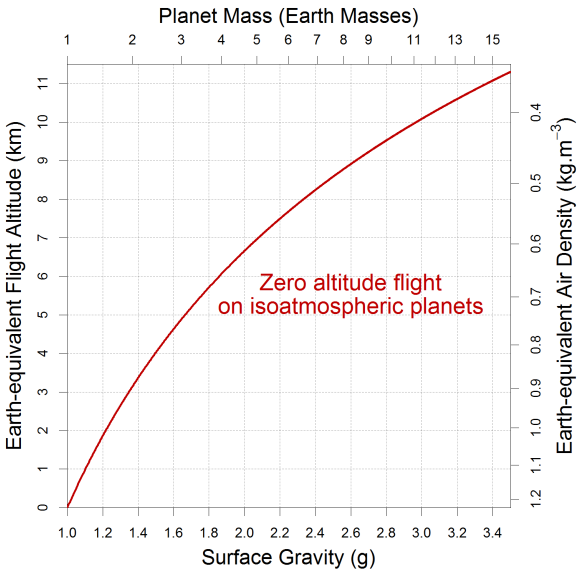

Fig. 4: For isoatmospheric planets the Earth-equivalent air density, ρ_{eq}, at the athleticism of zero-altitude flight, is inversely related to the surface gravity of a planet, $\rho_{eq} \propto 1/g_s$.

a whooper swan can be obtained by requiring that $\rho_s = \rho_{min}$. This implies that $\rho_w (M/M_\oplus)^{1.356} = \rho_\oplus (M/M_\oplus)^{0.904}$, and

$$M = M_\oplus \times \left(\frac{\rho_\oplus}{\rho_w}\right)^{1/0.452} \approx 6.86 M_\oplus. \qquad (13)$$

The surface gravity of this planet of 6.86 $M_\oplus$ would be $2.388g$. The maximum range and minimum power airspeeds of flying birds are expected to vary as $\rho^{-0.5}$ [12]. The surface air density of an isoatmospheric 6.86 Earth-mass planet would be $\sim 5.68\rho_\oplus$ so a typical airspeed of 21 m s^{-1} for a swan [24] might decline to 8.8 m s^{-1}, roughly the pace of an elite 400 m runner. In this same $2.385g$ environment, however, most people would struggle to walk at all and horses would be incapable of standing.

Since $P_f \propto m_b^{1+\alpha} \sqrt{g^3/\rho}$ and $\rho_s \propto M^{0.904}$, the flight power at zero altitude on isoatmospheric planets scales as $P_f \propto M^{0.226}$. Because $M/M_\oplus = (g_s/g_\oplus)^{1/0.452}$, it is apparent from (13) that $g_s/g_\oplus = \rho_\oplus/\rho_s$ for the limiting planet mass. Therefore, a particularly simple inverse relationship exists, $\rho_{eq} \propto 1/g_s$, allowing translation of the surface gravity of an isoatmospheric planet to the Earth-equivalent air density (and hence also equivalent maximum flight altitude via the International Standard Atmosphere). Results are presented in Fig. 4.

Might smaller birds be capable of flight on an isoatmospheric planet of 15 Earth masses? The surface air density would be $1.225\,(M/M_\oplus)^{0.904} = 14.17$ kg m^{-3}, lower than the minimum air density required by whooper swans for the same planet mass, $\rho_{min} = 20.18$ kg m^{-3}. Since $\chi \propto \rho_s^{-1/2}$, flight athleticism would have to be boosted by a factor of 1.1934. To achieve this, body mass could be reduced so that $m_b < m_w$ and flight would become feasible on a 15 Earth-mass planet if $m_b = m_w \times 1.1934^{-1/\alpha}$. Hence, flying animals of 0.42–3.8 kg or less (according to the value of α) may be capable of aerial locomotion on a 15 Earth-mass planet if they can match the flight athleticism of a whooper swan. Some ~88% of species have a body mass below 0.42 kg and ~99% have a body mass below 3.8 kg [3].

For an isoatmospheric 15-Earth mass planet one finds that $\sqrt{g_s^3/\rho_s} > 51.1$ m^3 s^{-3} kg$^{-1/2}$. On Earth this is equivalent to $\rho < 0.36$ kg m^{-3} or flight at altitudes $\gtrsim 11$ km. Even if smaller birds lack the athleticism of whooper swans, some may be able to fly in such rarefied air. The possibility could be investigated using a hypobaric wind tunnel operated at a comfortable flight temperature. Ruby-throated hummingbirds, which have a body mass of only 2–6 grams, can sustain hovering at densities down to 47% that of sea level air (0.576 kg m^{-3}) [29,30]. In forward flight, this species is likely to be capable of flying in yet more rarefied air. However,

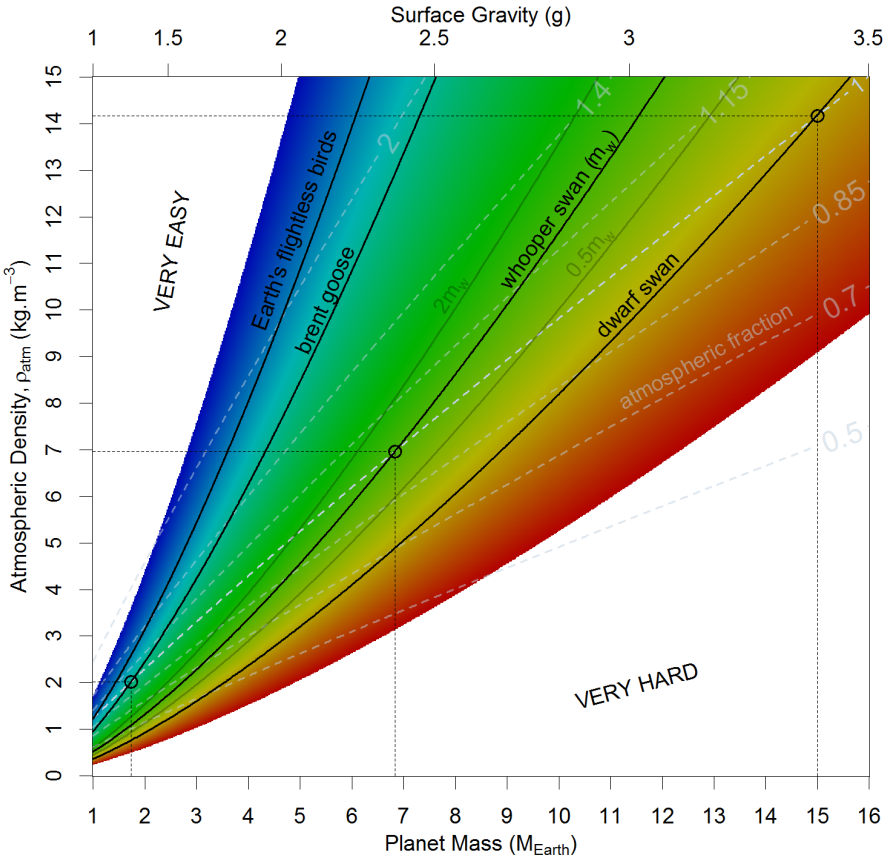

Fig. 5: Flight power (effort increases from blue to red) is a function of planet mass (or surface gravity) and atmospheric density. Conditions compatible with aerial locomotion lie upward of the solid contours. An 11 kg whooper swan appears capable of unassisted horizontal flight on isoatmospheric planets up to 6.86 $M_\oplus$. The influence of doubling or halving body mass relative to the whooper swan is shown for α=0.11. The trace marked *dwarf swan* corresponds to a hypothetical flying animal of the same flight athleticism as a whooper swan but of body mass 0.04–0.34m_w (corresponding to $0.054 < \alpha < 0.166$). Dashed contours represent atmospheric mass content relative to the Earth's fraction (862 parts per billion).

even then, due to its relatively small body mass, it is unlikely to challenge whooper swans for flight athleticism. The same argument applies also to flying insects.

Sufficient information has now been collected to describe circumstances compatible with environmentally unassisted circumplanetary flight in which buoyancy effects can be safely ignored. A planet would ideally occupy an orbit within the conventional circumstellar habitable zone [31] and, based upon the flight athleticism of whooper swans, the following criterion should also be satisfied:

$$\left(\frac{m_b}{m_w}\right)^{\alpha} \sqrt{\frac{g_s^3}{\rho_s}} \lesssim 42.8 \ \ \mathrm{m^3\,s^{-3}\,kg^{-1/2}}. \qquad (14)$$

By virtue of (12), an equivalent formulation involving only normalised mass terms is possible

$$\left(\frac{m_b}{m_w}\right)^{\alpha} \left(\frac{M}{M_\oplus}\right)^{0.226} \sqrt{\frac{M}{M_{atm}}} \lesssim 1646. \qquad (15)$$

Limitations in respiration or gas perfusion could potentially impinge upon the present analysis but oxygen delivery is not constrained in birds by the pulmonary system [23] and, in more inert atmospheres, flow-through breathing arrangements requiring little or no biomechanical effort can be imagined. Changes in atmospheric composition are likely over geological timescales [2]. Thus, it would ideally be useful to know whether an exoplanetary atmosphere has remained breathable and non-toxic for sufficient time to support the evolution of complex organisms.

Another factor which might well impact on these results is a change in atmospheric temperature, T_{atm}. The molar mass of the air, $M_{air} = 0.029\ \mathrm{kg\,mol^{-1}}$, the air temperature, T_{air}, and the universal gas constant, $R_{air} = 8,314\ \mathrm{N\,m\,mol^{-1}\,K^{-1}}$, obey the relationship $\beta = R_{air}T_{air}/M_{air}$. Since both γ and M_{atm} are linearly dependent on β, the value of $\sqrt{g_s^3/\rho_s}$ is proportional to $\sqrt{\beta}$. Since the value of β adopted here corresponds to an air temperature of 15°C, different atmospheric

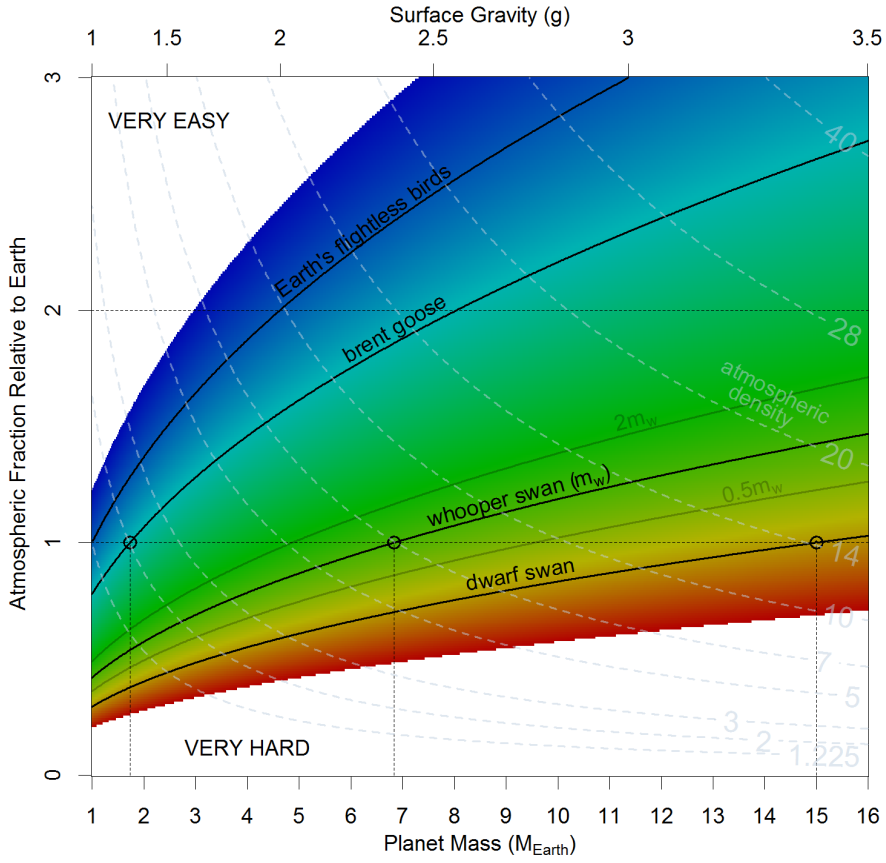

Fig. 6: Flight power according to planet mass and the atmospheric fraction relative to that of the Earth. Aerial locomotion is possible upwards of the solid contours. Dashed contours here represent the surface density of the atmosphere (kg m^{-3}), and correspond to isobars for the isothermal atmospheric model used here.

temperatures can be accommodated by applying a correction factor of $\sqrt{288.15/T_{atm}}$ to the right-hand sides of the inequalities (14) and (15).

The results of this analysis are presented graphically in Figs. 5 and 6. These limits are likely to be somewhat cautious since it is possible that, with determined effort, whooper swans may be capable of flying higher than the flock sighting at 8200 m. Although it has been conjectured that their initial ascent was aided by lee waves, such assistance would not have been present during the sea crossing from Iceland to Scotland [24]. Furthermore, this species regularly takes off in the dense air present at sea level which prohibits the evolution of larger wings that would tend to facilitate flight at extreme altitudes. Flying animals of extraterrestrial origin may not have been subjected to evolutionary pressures of this kind, particularly if their planets lack elevated land masses obstructing low altitude flight.

6 Discussion

Expressions (14) and (15) present criteria for aerial locomotion to be realistically possible in circumplanetary atmospheres. Comparisons of relative flight power under different environmental circumstances can utilise the expression $P_f \propto \sqrt{g_s^3/\rho_s}$. This predicts, for example, that flight in conditions resembling Saturn's moon Titan would be $\sim$ 23 times easier than at sea level on Earth. The wing-scaling exponent α has a small but positive value [15]. If this holds for a wide range of body masses then one can envisage animals flying in such conditions which are larger than any that have ever graced this planet. However, transport costs (or the energy/distance ratio), should approximately scale as $m_b^{0.7}$ during flight but only $m_b^{0.6}$ for running [1]. Above a certain body size, therefore, terrestrial locomotion would be energetically favoured to flight, though transit times might increase.

A primary finding is that, in the presence of a breathable atmosphere, winged animals of a body mass resembling the majority of the Earth's indigenous avian species could potentially evolve the ability to fly on isoatmospheric planets of at least 15 $M_\oplus$ ($g_s = 3.4g$). However, this work also highlights how even mildly reduced atmospheric fractions might potentially prohibit aerial locomotion. Novel techniques capable of remotely determining atmospheric composition, sur-

face atmospheric density and oceanic coverage could therefore be useful in augmenting future exoplanetary searches. Even worlds entirely covered in water could host flying animals. If in time the Earth were to become an ocean planet through continuing bombardment by comets and meteorites then seabirds could emulate penguins by mating, laying eggs and incubating them on floating icebergs.

That birds possess superb navigation skills has long been apparent but only recently have we appreciated that numerous species are adept problem-solvers [32] with an innate ability to fashion tools [33]. Eurasian Magpies (*Pica pica*) have demonstrated self-recognition when confronted with a mirror, a trait commonly associated with self-awareness [34]. Most birds are proficient hunters, potentially capable of stimulating the evolution of higher intelligence in land-based prey – such as our early mammalian ancestors. That cannot be said of insect-like creatures, which should in general cope more comfortably with higher gravitational fields due to the advantages of relatively small body masses and large area to volume ratios, facilitating respiration.

Flapping flight is a highly effective mode of locomotion for animals possessing sufficient athleticism. However, as aerial manoeuvres demand considerable coordination and spatiotemporal awareness, and body weight is critical, evolutionary pressures arise for efficient neurochemistry and neuroarchitecture. Volant organisms may well have played a pivotal role in shaping the Earth's natural history, enriching its biodiversity and accelerating the evolution of intelligent life. Avian species demonstrated considerable resilience in surviving the ecological catastrophe responsible for the extinction of most dinosaurs. In times of adversity, an ability to swiftly and efficiently relocate over planetary distances and flexibly forage on both land and sea may assist the propagation of flying animals over geological∼stellar timescales. Accurate determination of whether circumplanetary flight is possible should not be overlooked if future missions to extrasolar worlds are intent on maximising the chances of encountering complex lifeforms and, perhaps, even extraterrestrial civilisations of comparable sophistication to our own.

Submitted on May 13, 2016 / Accepted on May 17, 2016

References

1. Schmidt-Nielsen K. *Science*, 1972, v. 177 (4045), 222–228.
2. Dudley R. *Ann. Rev. Physiol.*, 2000, v. 62 (1), 135–155.
3. Codron D., Carbone C. & Clauss M. *PLoS One*, 2013, v. 8 (10), e77110.
4. Egevang C., Stenhouse I. J., Phillips R. A., Petersen A., Fox J. W. & Silk J. R. *Proceedings of the National Academy of Sciences*, 2010, v. 107 (5), 2078–2081.
5. Spivey R. J., Bishop C. M., Butler P. J. *et al. Science*, 2015, v. 347 (6219), 250–254.
6. Gatesy S. M. & Dial K. P. *Evolution*, 1996, v. 50 (1), 331–340.
7. Pearson R. G. The Avian Brain. Academic Press, 1972, ISBN 9780125480505.
8. Marcy G. W., Isaacson H., Howard A. W., Rowe J. F., Jenkins J. M., Bryson S. T., Christiansen J. *et al. Astrophys. J. Suppl. Ser.*, 2014, v. 210 (2), 20.
9. Cassan A., Kubas D., Beaulieu J. P., Dominik M., Horne K., Greenhill J., Pietrzyski G. *et al. Nature*, 2012, v. 481 (7380), 167–169.
10. Petigura E. A., Howard A. W. & Marcy G. W. *P.N.A.S.*, 2013, v. 110 (48), 19273–19278.
11. Tobalske B. W., Hedrick T. L., Dial K. P. & Biewener A. A. *Nature*, 2003, v. 421 (6921), 363–366.
12. Pennycuick C. J. Modelling the Flying Bird (Vol. 5). Elsevier, 2008, ISBN 9780123742995.
13. Altshuler D. L. & Dudley R. *Int. & Comp. Biol.*, 2006, v. 46 (1), 62–71.
14. Rayner J. *Mathematical Methods Applied Science*, 2001, v. 24 (17–18), 1485–1514.
15. Hedenstrom A. *J. Comp. Physiol. A*, 2008, v. 194 (7), 685–691.
16. Scott G. R. *J. Exp. Biol.*, 2011, v. 214 (15), 2455–2462.
17. Black C. R., Tenney S. M., *Respir. Physiol.*, 1980, v. 39, 217–239.
18. Hawkes L. A., Balachandran S., Batbayar et al. *Proc. Roy. Soc. B*, 2012, v. 280 (1750).
19. Gudmundsson G. A., Benvenuti S., Alerstam T., Papi F. & Lilliendahl K. *Proc. Roy. Soc. B*, 1995, v. 261 (1360), 73–79.
20. Blum A. Annapurna: A Woman's Place. Sierra Club Books, 1980, ISBN 9781578050222.
21. Tucker V. A. *Journal of Experimental Biology*, 1968, v. 48 (1), 55–66.
22. Fedde M. R., Faraci F. M., Kilgore Jr D. L., Cardinet III G. H. & Chatterjee A. Circulation, Respiration, and Metabolism. Springer Berlin Heidelberg, 1985, 149–163.
23. Tobalske B. W. & Dial K. P. *J. Exp. Biol.*, 2000, v. 203 (21), 3319–3332.
24. Pennycuick C. J., Einarsson O., Bradbury T. A. M. & Owen M. *J. Avian Biol.*, 1996, v. 27 (2), 118–134.
25. Liechti F. & Schaller E. *Naturwissenschaften*, 1999, v. 86 (11), 549–551.
26. Sotin C., Grasset O. & Mocquet A. *Icarus*, 2007, v. 191 (1), 337–351.
27. Spivey R. J. *J. Modern Physics*, 2013, v. 4, 20–47.
28. Spivey R. J. *Physics Essays*, 2015, v. 28 (2), 254–264.
29. Chai P. & Dudley R. *Nature*, 1995, v. 377 (6551), 722–725.
30. Chai P. & Dudley R. *J. Exp. Biol.*, 1995, v. 199 (10), 2285–2295.
31. Kasting J. F., Whitmire D. P. & Reynolds R. T. *Icarus*, 1993, v. 101 (1), 108–128.
32. Emery N. J. *Phil. Trans. Roy. Soc. B*, 2006, v. 361 (1465), 23–43.
33. Hunt G. R. *Nature*, 1996, v. 379 (6562), 249–251.
34. Prior H., Schwarz A. & Gunturkun O. *PLoS Biology*, 2008, v. 6 (8), e202.

The Structure of the Photon in Complex Vector Space

Kundeti Muralidhar

Physics Department, National Defence Academy, Khadakwasla, Pune-411023, India. E-mail: kundetimuralidhar@gmail.com

Considering the complex vector electromagnetic field, the energy of the photon is expressed as an even multivector consisting of a scalar kinetic energy part and a bivector rotational energy part. Since any even multivector can be expressed as a rotor representing internal rotations, the electromagnetic energy even multivector represents internal complex rotations. It has been shown that the spin angular momentum is the generator of rotations in the plane normal to the propagation direction and the orbital angular momentum is the generator of rotations in a plane normal to the spin plane. The internal structure of the photon may be visualized as a superposition of electromagnetic field flow or rotation in two normal orientations in complex vector space. The cause of such complex rotations is attributed to the presence of electromagnetic zeropoint field.

1 Introduction

Even after the photon inception into the field of physics over a century ago, the obscurity in understanding the photon structure persists. The concept of the photon, the energy quanta of electromagnetic radiation, was introduced by Planck in the blackbody radiation formula and Einstein in the explanation of the photoelectric effect. The photon is normally considered as a massless bundle of electromagnetic energy and the photon momentum is defined as the ratio between the energy of the photon and the velocity of light. It is well known from Maxwell's theory that electromagnetic radiation carries both energy and momentum [1]. The linear momentum density is given by the Poynting vector $\mathbf{E} \times \mathbf{B}$ and the angular momentum is the cross product of the Poynting vector with the position vector. Poynting suggested that circularly polarized light must contain angular momentum and showed it as the ratio between the free energy per unit volume and the angular frequency. In 1936, Beth [2] first measured the angular momentum of light from the inference that circularly polarized light should exert torque on a birefringent plate and that the ratio between angular momentum J and linear momentum P was found to be $\lambda/2\pi$, where λ is the wavelength of light. The measured angular momentum agreed in spin magnitude with that predicted by both wave mechanics and quantum mechanics. The Beth angular momentum is in general considered as the photon spin angular momentum.

The energy momentum tensor of the electromagnetic field $T^{\mu\nu}$ is not generally symmetric. By adding a divergence term $\partial_\mu U^{\mu\alpha\nu}$ to $T^{\mu\nu}$, one can construct a symmetric energy momentum tensor $\Theta^{\mu\nu}$ which is normally known as the Belinfante energy momentum tensor [3]. The tensor $U^{\mu\alpha\nu}$ is asymmetric in the last two indices. The symmetric energy momentum tensor satisfies the conservation law $\partial_\mu \Theta^{\mu\nu} = 0$. The advantage of the symmetric energy momentum tensor is that the angular momentum calculated from Θ^{k0} is a conserved quantity. Belinfante established the fact that the spin could be regarded as a circulating flow of energy and this idea was well explained by Ohanian [4]. In an infinite plane wave, the electric and magnetic field vectors are perpendicular to the propagation direction. In a finite transverse extent, the field lines are closed loops and represent circulating energy flow and imply the existence of angular momentum whose orientation is in the plane of circulation and it is the spin angular momentum. Further, as the electromagnetic waves propagate, the energy also flows along the direction of propagation. The translational energy flow implies the existence of additional orbital angular momentum. The magnetic field vector can be expressed as the curl of a vector potential $\mathbf{A}$ and the angular momentum density becomes $\mathbf{E} \times \mathbf{A}$. A close inspection shows that the total angular momentum has two components: one the spin angular momentum associated with the polarization and the other the orbital angular momentum associated with the spatial distribution [1]. The total angular momentum J can be split into a spin angular momentum S and an orbital angular momentum L [5]

$$J = \frac{1}{4\pi} \int \mathbf{E} \times \mathbf{A}\, d^3 r + \frac{1}{4\pi} \int \mathbf{E}^n (\mathbf{r} \times \nabla) \mathbf{A}^n\, d^3 r\,. \qquad (1)$$

The first term on the right is dependent on polarization and hence it is called spin angular momentum S and the second term is independent of polarization and depends on spatial distribution and identified with orbital angular momentum L. It has been argued that the photon angular momentum cannot be separated into a spin part and an orbital part in a gauge invariant way and the paradox was a subject for several papers and in standard textbooks for the past few decades [6].

In recent times the definitions of these angular momenta raised certain controversy. In all these definitions the angular momentum is defined as a vector product containing the position coordinate. The decomposition of total angular momentum of the photon into spin and orbital parts basically involves how we split the vector potential in a gauge invariant way and it has been studied by several authors and a detailed discussion is given in the review article by Leader and Lorce [7]. The absence of any rest frame for the photon suggests that the total angular momentum is observable but not separately

as spin angular momentum and orbital angular momentum. Though this separation is normally considered to be unphysical and not observable, Van Enk and Nienhuis [8] argued that both spin and orbital angular momenta are separately measurable quantities and gauge invariant. The gauge invariant spin and angular momentum parts are expressed as

$$J = \frac{1}{4\pi}\int \mathbf{E}\times\mathbf{A}\, d^3r + \frac{1}{4\pi}\int \mathbf{r}\times\mathbf{E}^n\nabla\mathbf{A}^n\, d^3r\,. \qquad (2)$$

In this expression $\mathbf{A} = \mathbf{A}_\perp$ and therefore both terms are gauge invariant. The canonical expression $\mathbf{E}^n\nabla\mathbf{A}^n$ gives pure mechanical momentum which is responsible for the orbital angular momentum of a photon. The azimuthal flow of electromagnetic field is given by $\mathbf{E}^n\nabla\mathbf{A}^n$ which is half of $\mathbf{E}\times\mathbf{B}$ and the other half is spin flow [9]. In an analogous way, in quantum chromodynamics, the gluon angular momentum can be decomposed into a spin part and an angular momentum part which plays an important role in understanding nucleon structure. Recently, Chen *et al.* [10] decomposed the gauge potential into pure and physical parts: $\mathbf{A} = \mathbf{A}_{pure} + \mathbf{A}_{phys}$, the pure part is related to gauge invariance and the physical part is related to physical degrees of freedom. In the decomposition by Wakamatsu *et al.* [11], the orbital angular momentum is defined similar to a classical expression $\mathbf{r}\times P_{kin}$, where $P_{kin} = -\frac{1}{4\pi}\int \mathbf{A}_{phys}\times\mathbf{E}\, d^3r$ and in this decomposition each term is gauge invariant and observable. Further studies by several authors revealed the fact that there could be infinitely many different ways to perform such decomposition in a gauge invariant way [7, 12]. In Beth's experiment, actually the spin angular momentum was measured. The measurement of orbital angular momentum has been performed in recent times. The amplitude of a Laguerre-Gaussian mode of light wave has an azimuthal angular dependence of $\exp(-il\phi)$, where l is the azimuthal mode index. The ratio between the angular momentum to the energy is $1/\omega$ or $L = l\,(\mathbf{E}\times\mathbf{B})/\omega$ and for Laguerre-Gaussian laser mode, it has been shown that the angular momentum is equal to $l\hbar$ and the total angular momentum of the whole light beam is $(l+\sigma_z)\hbar$, where σ_z is a unit vector along the direction of propagation [13]. The measurement of orbital angular momentum was reported by several authors [14–16].

Another important aspect of the photon is its internal zitterbewegung motion. It is well known from the first observation of Schrödinger [17] that a Dirac electron possesses zitterbewegung motion which is the oscillatory motion of the electron with very high frequency $\omega = 2mc^2/\hbar$ with internal velocity equal to the velocity of light. Such internal motion arises because of the classical electromagnetic fluctuating zeropoint field present throughout space [18]. The spin angular momentum of the electron is identified as the zeropoint angular momentum [18, 19]. On the basis of electron internal oscillations, classical models of electron were developed [20–22]. It is quite interesting that such zitterbewegung motion for the photon was derived from the relativistic Schrödinger like equation of the photon by Kobe [23]. It has been proved that the photon velocity contains parallel and perpendicular components with respect to the direction of propagation. The time dependent perpendicular component of velocity rotates about the direction of propagation with an angular frequency ω equal to the frequency of the electromagnetic wave. The finite special extension of internal rotation is equal to the reduced wavelength. The photon spin is then identified as the internal angular momentum due to zitterbewegung. Considering internal dynamical variables in the configuration space the zitterbewegung is attributed to the normal component of velocity vector oscillations about the particle centre [24]. In the quantum field theory, it has been shown that the zitterbewegung of a photon is attributed to the virtual transition process corresponding to the continuous creation and annihilation of virtual pairs of elementary excitations [25, 26]. Recently, Zhang [27] proposed that zitterbewegung of the photons may appear near the Dirac point in a two dimensional photonic crystal. In the case of an electron, the spin angular momentum is an intrinsic property. In the same way both spin and orbital momenta of the photon are intrinsic in nature [28, 29]. Thus one can anticipate that the photon is also having an internal spin structure described by the internal oscillations or rotations.

One of the most important applications of the photon angular momentum lies in the exploitation of the photon spin and angular momentum states for quantum computation and quantum information processing [30]. Superposition of polarization states can be used to construct qubits and transmit information. A standard approach to visualise the transformation of qubits is provided by the Poincaré sphere representation. Generally, any completely polarized state can be described as a linear superposition of spin states and corresponds to a point on the surface of a unit sphere. Analogous representation of orbital angular momentum states of the photon was introduced by Padgett and Courtial [31] and Agarwal [32]. Quantum entanglement of states is a consequence of quantum non-locality. The entanglement involving the spatial modes of electromagnetic field carrying orbital angular momentum was studied by Mair *et al.* [33] and Franke-Arnold *et al.* [34]. The phase dependence of angular momentum may provide multi-dimensional entangled states which are of considerable interest in the field of quantum information.

In vector algebra, the angular momentum is defined by a cross product of position and momentum vectors and identified as a vector normal to the plane containing the position and momentum vectors. However, the angular momentum is basically a planar quantity and better defined as a bivector in a plane [35]. Note that the cross product cannot be defined in a plane. In the case of the electron, the classical internal bivector spin was obtained from the multivector valued Lagrangian by Barut and Zhangi [20]. It has been shown that the particle executes internal complex rotations by absorbing zeropoint field and the angular momentum of these internal rotations

is identified as the bivector spin of the particle [36]. In an analogous way, the photon spin is a bivector quantity. Therefore, the photon spin may be visualised as a bivector product between an internal finite extension of the photon and an internal momentum. Similarly, one can visualise the orbital angular momentum of the photon as a bivector. Considering the electromagnetic field as a complex vector, it is possible to express the set of Maxwell's equations into a single form. The basics of complex vector algebra have been discussed in detail previously in references [37, 38].

Recently, the nature of the photon was discussed at length by several authors in the book edited by Roychoudhuri *et al.* [39]. The main views of understanding the nature or the structure of the photon are as follows. Einstein viewed the photon as a singular point which is surrounded by electromagnetic fields. In quantum electrodynamics, the photon is introduced as a unit of excitation associated with the quantised mode of the radiation field and it is associated with precise momentum, energy and polarization. In another view, the photon is interpreted as neither a quantum nor a wave but it can be a meson which produces off other hadronic matter and attains physical status. Photons are just fluctuations of random field or wave packets in the form of needles of radiation superimposed in the zeropoint field. However, understanding the photon structure still remains an open question.

The aim of this article is to explore the structure of the photon in complex vector space. To understand the structure of the photon, the electromagnetic field is expressed as a complex vector and the total energy momentum even multivector is developed in section 2. Section 3 deals with the internal angular momentum structure of the photon and conclusions are presented in section 4. Throughout this article, Lorentz–Heaviside units are used, i.e. $\epsilon_0 = \mu_0 = 1$ and energy terms are divided by 4π and conveniently we choose $c = 1$ [1]. However, for clarification sake in some places c is reintroduced.

2 Energy momentum multivector of the electromagnetic field

In the complex vector formalism, we express the electric field as a vector $\mathbf{E}$ and the magnetic field as a bivector $\mathbf{iB}$ and the electromagnetic field F is expressed as a complex vector [37, 38]

$$F = \frac{1}{2}(\mathbf{E} + \mathbf{iB}) . \tag{3}$$

Here, $\mathbf{i}$ is a pseudoscalar in geometric algebra of three dimensions [40], it commutes with all elements of the algebra and $\mathbf{i}^2 = -1$. A reversion operation changes the order of vectors and is indicated by an overbar

$$\bar{F} = \frac{1}{2}(\mathbf{E} - \mathbf{iB}) . \tag{4}$$

Now, the product $\bar{F}F$ is written as

$$\bar{F}F = \frac{1}{4}(E^2 + B^2) + \frac{1}{2}(\mathbf{E} \wedge \mathbf{iB}) . \tag{5}$$

Similarly, we find

$$F\bar{F} = \frac{1}{4}(E^2 + B^2) - \frac{1}{2}(\mathbf{E} \wedge \mathbf{iB}) . \tag{6}$$

The energy density of the electromagnetic field can be obtained from the scalar product

$$\bar{F} \cdot F = \frac{1}{2}(\bar{F}F + F\bar{F}) = \frac{1}{4}(E^2 + B^2) . \tag{7}$$

Further, the product $\bar{F} \wedge F$ gives a vector of the form

$$\mathbf{p} = -\frac{1}{c}\bar{F} \wedge F = -\frac{1}{2c}(\bar{F}F - F\bar{F}) = -\frac{1}{2c}(\mathbf{E} \wedge \mathbf{iB}) , \tag{8}$$

and the dual of $\mathbf{p}$ is expressed as

$$\mathbf{ip} = \frac{1}{2c}(\mathbf{E} \wedge \mathbf{B}) . \tag{9}$$

From the above expression, one can express the energy density of internal electromagnetic flux flow in the bivector plane normal to the propagation direction

$$\mathbf{ip}c = \frac{1}{2}(\mathbf{E} \wedge \mathbf{B}) . \tag{10}$$

This energy density of the photon can be identified as the rotational energy density. However, the energy density obtained in (7) represents the energy density of the photon as it propagates and it may be treated as the kinetic energy density of the photon. An even multivector is a sum of a vector and a bivector. The energy terms in (7) and (10) combine to give the total energy of the photon in even multivector form

$$\mathcal{E} = \frac{1}{4\pi}\int \frac{E^2 + B^2}{4} d^3r + \frac{1}{4\pi}\int \frac{\mathbf{E} \wedge \mathbf{B}}{2} d^3r = \mathcal{E}_{kin} + \mathcal{E}_{rot} . \tag{11}$$

The scalar part shows the flow of energy in the direction of propagation which can be identified as the kinetic part of energy $\mathcal{E}_{kin}$ and the bivector part can be identified as the rotational energy $\mathcal{E}_{rot}$ representing circulation of electromagnetic energy in a plane normal to the direction of propagation. In general, twice the kinetic energy is treated as the electromagnetic energy per unit volume and it is the energy of the photon. Since the energy of a photon is expressed as momentum times its velocity, we define kinetic momentum of a photon as $\mathbf{p}_k = \mathcal{E}_{kin}/\mathbf{v}$, where the velocity $\mathbf{v} = \mathbf{n}c$ and $\mathbf{n}$ is a unit vector along the direction of propagation. Introducing an internal velocity $\mathbf{u}$ satisfying the condition $\mathbf{u} \cdot \mathbf{v} = 0$ and $|\mathbf{u}| = c$, the internal momentum representing the rotational flux flow can be defined as $\mathbf{p}_r = \mathcal{E}_{rot}/\mathbf{u}$. From these definitions generalised photon velocity and momentum complex vectors can be constructed as

$$U = \mathbf{v} + \mathbf{iu} , \tag{12}$$

$$P = \mathbf{p}_k + \mathbf{i}\mathbf{p}_r \,. \tag{13}$$

A reversion operation on P gives $\bar{P} = \mathbf{p}_k - \mathbf{i}\mathbf{p}_r$. Since the magnitudes $|\mathbf{p}_k|$ and $|\mathbf{p}_r|$ are equal, we have $P^2 = \bar{P}^2 = 0$. Therefore, the complex vector P is a complex null vector which represents the lightlike nature of the photon. Similarly, the complex velocity vector is also a complex null vector. Now, the total energy of the photon is expressed as

$$\mathcal{E} = \mathbf{p}_k \cdot \mathbf{v} + \mathbf{p}_r \wedge \mathbf{u} \,. \tag{14}$$

The even multivector form given in the above equation can be compared to the symmetric energy momentum tensor $\Theta^{\mu\nu}$ with the identification of the scalar part with Θ^{00} and bivector part with Θ^{ij} . In three dimensions, the property of an even multivector is that it represents rotations in the bivector plane [35]. Then, the energy multivector can be expressed as a rotor with angular frequency ω

$$\mathcal{E} = \mathcal{E}_0 \, e^{\hat{J}\omega t} \,, \tag{15}$$

where $\hat{J}$ is a unit bivector in the plane normal to the propagation direction. This relation shows that the photon contains internal complex rotations and these rotations are analogous to the internal complex rotations or zitterbewegung of the electron. The cause of these internal rotations is attributed to the fluctuations of the zeropoint field [38]. In (15) the internal rotation represents the clockwise or right-handed rotation. A reversion operation on $\mathcal{E}$ gives

$$\bar{\mathcal{E}} = \mathcal{E}_0 \, e^{-\hat{J}\omega t} \,. \tag{16}$$

In this case, the internal rotation represents counterclockwise or left-handed rotation. The frequency of internal rotation is the rate per unit energy flux flow within the photon

$$\Omega = -\hat{J}\omega t = -\frac{1}{\mathcal{E}} \frac{d\mathcal{E}}{dt} \,. \tag{17}$$

Here, the frequency of internal rotation represents the counterclockwise direction. The internal complex rotations suggest that there exists an internal complex structure of the photon.

3 Internal structure of the photon

In general, the internal complex rotations represent the angular momentum of the photon. The angular momentum of a photon is defined as the ratio between the rotational energy of the photon and the frequency of internal rotation. Since the energy of the photon is a sum of kinetic and rotational energy components, we expect that the angular momentum of the photon contains two parts: one corresponding to the rotational flow of energy and the other to the translational flow of energy. According to the definition given in (2), the spin angular momentum bivector is in the orientation of the plane $\mathbf{A} \wedge \mathbf{E}$ which is a plane normal to the propagation direction. Let us consider a set of orthogonal unit vectors $\{\sigma_k; k = 1, 2, 3\}$ along x, y and z axes. If we choose the propagation direction along the z-axis, then the unit bivector along the spin orientation is $\mathbf{i}\sigma_3$. To understand the orientation of spin and orbital angular momenta, let us consider circularly polarized light waves propagating along the z-direction and the waves have finite extent in the $x-$ and y-directions. The propagating wave has cylindrical symmetry about the z-axis. The energy of the wave can be visualised as a sum of circulating energy flow in the x-y plane and a translational energy flow in the z-direction. In the case of circularly polarized light the vector potential $\mathbf{A}$ contains only two components

$$\mathbf{A} = \frac{\mathcal{E}_0}{\omega} \left[\sigma_1 cos(\mathbf{k}\cdot\mathbf{r} - \omega t) + \sigma_2 sin(\mathbf{k}\cdot\mathbf{r} - \omega t)\right] \,. \tag{18}$$

Here, $\mathbf{k}$ is the wave vector. The three vectors $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{A}$ rotate in a plane normal to the propagation direction. Differentiation of (18) with respect to time gives the electric field vector

$$\mathbf{E} = \mathcal{E}_0 \left[-\sigma_1 sin(\mathbf{k}\cdot\mathbf{r} - \omega t) + \sigma_2 cos(\mathbf{k}\cdot\mathbf{r} - \omega t)\right] \,. \tag{19}$$

Then, the bivector product $\mathbf{A} \wedge \mathbf{E}$ becomes

$$\mathbf{A} \wedge \mathbf{E} = \mathbf{i}\sigma_3 \frac{\mathcal{E}_0^2}{\omega} \,, \tag{20}$$

where $\sigma_1\sigma_2 = \mathbf{i}\sigma_3$. The spin angular momentum of electromagnetic field or the photon is expressed as

$$S = \mathbf{i}\sigma_3 \frac{1}{4\pi} \int \frac{\mathcal{E}_0^2}{\omega} d^3r = \mathbf{i}\sigma_3 \hbar \,, \tag{21}$$

where the energy density of the electromagnetic wave is normalized so that the energy is one quantum. Normally, because of the fact $\sigma_3 \wedge \mathbf{k} = 0$, the z-component of angular momentum goes to zero but not the other components of the orbital angular momentum. From the second term on the right of (2), the angular momentum density is expressed as a sum of two terms

$$\mathbf{r} \wedge \mathbf{E}^n \nabla \mathbf{A}^n = \mathbf{r} \wedge \mathbf{E}^x \nabla \mathbf{A}^x + \mathbf{r} \wedge \mathbf{E}^y \nabla \mathbf{A}^y \,. \tag{22}$$

Substituting individual components of $\mathbf{E}$ and $\nabla\mathbf{A}$ in the above equation, we find the orbital angular momentum density

$$\mathbf{r} \wedge \mathbf{E}^n \nabla \mathbf{A}^n = \frac{\mathcal{E}_0^2}{\omega} \mathbf{r} \wedge \mathbf{k} \,. \tag{23}$$

Then the orbital angular momentum of the photon is expressed as

$$L = \mathbf{r} \wedge \mathbf{k} \frac{1}{4\pi} \int \frac{\mathcal{E}_0^2}{\omega} d^3r \,. \tag{24}$$

In the above equation, the vector $\mathbf{r}$ is restricted to the plane $\mathbf{i}\sigma_3$ and contains only x and y components. If the magnitude of $\mathbf{r}$ is equal to the reduced wavelength, then the product

$|\mathbf{r}||\mathbf{k}| = 1$ for circularly polarized light. The orbital angular momentum is now expressed as

$$L = \hat{\mathbf{r}} \wedge \sigma_3 \frac{1}{4\pi} \int \frac{\mathcal{E}_0^2}{\omega} d^3r = \mathbf{im}\hbar\,, \tag{25}$$

where the unit vector $\hat{\mathbf{r}}$, in an arbitrary direction, lies in the plane $\mathbf{i}\sigma_3$, the unit vector $\mathbf{m}$ is chosen normal to the orientation of $\hat{\mathbf{r}} \wedge \sigma_3$ and the integral term in (25) represents the ratio between the energy of the photon and the frequency. Thus the orientation of the orbital angular momentum is always normal to the orientation of spin angular momentum in a photon. In the case the photon is propagating in an arbitrary direction say $\mathbf{n}$ then from the above analysis, the spin angular momentum and orbital angular momentum are expressed as

$$S = \mathbf{in}\hbar\,, \tag{26}$$

$$L = \mathbf{im}\hbar\,. \tag{27}$$

The vectors $\mathbf{n}$ and $\mathbf{m}$ satisfy the condition $\mathbf{n} \cdot \mathbf{m} = 0$ and the vector $\mathbf{m}$ lies in the plane of the unit bivector $\mathbf{in}$. Since, the direction of unit vector $\mathbf{m}$ or the orientation of the plane $\mathbf{im}$ is arbitrary, the rotation of $\mathbf{m}$ is expressed by the relation $\mathbf{m}' = \bar{R}\mathbf{m}R$. Here, $R = e^{\mathbf{in}\phi/2}$ is a rotor and in this way the orbital angular momentum depends on the angle ϕ. The spin angular momentum describes the intrinsic angular momentum of a photon and commutes with the generator of translation $\mathbf{n}|\mathbf{k}|$. The spin angular momentum causes the complex vector field F to rotate in the $\mathbf{E} \wedge \mathbf{B}$ plane without changing the direction of propagation vector $\mathbf{k}$. The photon spin is the generator of rotations in the plane normal to the propagation direction. Whereas, the orbital angular momentum causes the plane having orientation defined by the bivector $\mathbf{r} \wedge \mathbf{k}$ to rotate without changing the direction of the vector $\mathbf{k}$ and the orientation of the plane $\mathbf{E} \wedge \mathbf{B}$. The orbital angular momentum does not commute with the generator of translation. The photon orbital angular momentum is the generator of rotations in a plane normal to the spin plane. Thus one can conclude that both the spin and orbital angular momenta of a photon are intrinsic. The intrinsic nature of orbital angular momentum was discussed by Berry [28]. Further, Allen and Padgett [29] argued that the spin and the orbital angular momenta are intrinsic in nature in the case when the transverse momentum is zero for the helical wave fronts. The spin and orbital angular momenta of the photon are fundamental quantities and produce complex rotations in space and such rotations are actually produced by the fluctuating zeropoint fields present throughout space [38, 41]. The internal complex rotations are not only limited to the rotations pertaining to the plane of spin angular momentum but also exists in the plane of orbital angular momentum. In the Laguerre-Gaussian modes of laser beams it has been shown explicitly in the quantum mechanical approach that the orbital angular momentum of light beams resembles the angular momentum of the harmonic oscillator [42].

4 Conclusions

The electromagnetic field per unit volume is represented by an energy momentum even multivector and expressed as a sum of scalar and bivector components, and we identify the scalar part as the kinetic part which shows the flow of energy in the direction of propagation and the bivector part as the rotational energy flow in the plane normal to the direction of propagation over a finite extent. The even multivector form of energy shows that there exist internal complex rotations of the electromagnetic field. The cause of these internal rotations is attributed to the fluctuations of the zeropoint field. In general, the internal complex rotations represent the angular momentum of the photon. The angular momentum of the photon is defined as the ratio between the rotational energy of the photon and the angular frequency of rotation. The spin angular momentum bivector represents a plane normal to the propagation direction. We find that the orientation of orbital angular momentum is always normal to the orientation of spin angular momentum in a photon. The photon spin is the generator of rotations in the plane normal to the propagation direction. The photon orbital angular momentum is the generator of rotations in a plane normal to the spin plane. Thus, one can conclude that both spin and orbital angular momenta of a photon are intrinsic in nature. The internal structure of the photon may be visualized as the superposition of electromagnetic field flow or rotation in two normal orientations in complex vector space. Because of the formal similarity between gluons and photons, the conclusions obtained here may be extended to the gluon structure.

Submitted on May 23, 2016 / Accepted on May 30, 2016

References

1. Jackson J. D. Classical Electrodynamics. Wiley Eastern Limited, New Delhi, 1978.
2. Beth R. A. Mechanical Detection and Measurement of the Angular Momentum of Light. *Phys. Rev.*, 1936, v. 50, 115.
3. Belinfante F. J. On the spin angular momentum of mesons. *Physica*, 1939, v. 6, 887–898.
4. Ohanian H. C. What is spin? *Am. J. Phys.*, 1986, v. 54, 500.
5. Cohen-Tannoudji C., Dupont-Roc J., Grynberg G. Photons and Atoms. John Wiley & Sons, New York, 1989.
6. Jauch J. M., Rohlich F. The Theory of Photons and Electrons. Springer-Verlag, Berlin, 1976.
7. Leader E., Lorce C. The angular momentum controversy: What's it all about and does it matter? *Physics Reports*, 2014, v. 541, 163–248.
8. Van Enk, Nienhuis G. Spin and Orbital Angular Momentum of Photons. *Euro. Phys. Lett.*, 1994, v. 25, 497.
9. Chen X. S., Chen G. B., Sun W. M., Wang F. Testing the correctness of the Poynting vector $\mathbf{E} \times \mathbf{B}$ as the momentum density of gauge fields. arXiv: hep-ph/0710.1427v1.
10. Chen X. S., Lu X. F., Sun W. M., Wang F., Goldman T. Spin and Orbital Angular Momentum in Gauge Theories: Nucleon Spin Structure and Multipole Radiation Revisited. *Phys. Rev. Lett.*, 2008, v. 100, 232002.
11. Wakamatsu M. Gauge-invariant decomposition of nucleon spin. *Phys. Rev. D.*, 2010, v. 81, 114010.

12. Sun W. M. Physical Decomposition of Photon Angular Momentum. arXiv: quant-ph/1407.2305v1.

13. Allen L., Beijersbergen M. W., Spreeuw R. J. C., Woerdman J. P. Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes. *Phys. Rev. A*, 1992, v. 45, 8185.

14. Leach J., Padgett M. J., Barnett S. M., Franke-Arnold S., Courtial J. Measuring the orbital angular momentum of a single photon. *Phys. Rev. Lett.*, 2002, v. 88, 257901.

15. Allen L., Barnett S. M., Padgett M. J. Optical Angular Momentum. IOP Publishing, Bristol and Philadelphia, 2003.

16. Padgett M., Courtial J., Allen L. Light's Orbital Angular Momentum. *Physics Today*, May 2004, 35.

17. Barut A. O., Bracken A. J. Zitterbewegung and the internal geometry of electron. *Phys. Rev. D*, 1981, v. 23, 2454.

18. Hestenes D. Spin and uncertainty in the interpretation of quantum mechanics. *Am. J. Phys.*, 1979, v. 47, 399–415.

19. Hestenes D. Zitterbewegung in Quantum Mechanics. *Found. Phys.*, 2010, v. 40, 1–54.

20. Barut A. O., Zanghi A. J. Classical Model of the Dirac Electron. *Phys. Rev. Lett.*, 1984, v. 52, 2009–2012.

21. Hovarthy P. A. Mathisson's spinning electron: Noncommutative mechanics and exotic Galilean symmetry, 66 years ago. *Acta. Phys. Pol. B*, 2003, v. 34, 2611–2622.

22. Vaz J., Rodriguez W. A., Zitterbewegung and the electromagnetic field of the electron. *Phys. Lett. B*, 1993, v. 319, 243.

23. Kobe D. H. Zitterbewegung of a photon. *Phys. Lett. A*, 1999, v. 253, 7–11.

24. Unal N. A Simple Model of Classical Zitterbewegung: Photon Wave Function. *Found. Phys.*, 1997, v. 27, 731.

25. Wang Z. Y., Xiong C. D. Zitterbewegung by quantum field-theory considerations. *Phys. Rev. A*, 2008, v. 77, 045402.

26. Wang Z. Y., Xiong C. D., Qiu Q. Photon wave function and Zitterbewegung. *Phys. Rev, A*, 2009, v. 80, 032118.

27. Zhang X. Observing Zitterbewegung for Photon near the Dirac point of a Two-Dimensional Photonic Crystal. *Phys. Rev. Lett.*, 2008, v. 100, 113903.

28. Berry M. V. Paraxial beams of spinning light. In: Allen S., Barnett S. M., Padgett, M. J., eds. Optical Angular Momentum. IOP Publishing, Bristol, Philadelphia, 2003.

29. Allen L., Padgett M. J. The Poynting vector in Laguerre-Gaussian beams and the interpretation of their angular momentum density. *Opt. Commun.*, 2000, v. 184, 6771.

30. Calvo G. F., Picón A., Bagan E. Quantum field theory of photons with orbital angular momentum. *Phys. Rev. A*, 2006, v. 73, 01385.

31. Padgett M. J., Courtial J. Poincare-sphere equivalent for light beams containing orbital angular momentum. *Opt. Lett.*, 1999, v. 24, 430.

32. Agarwal G. S. SU(2) structure of the Poincaré sphere for light beams with orbital angular momentum. *J. Opt. Soc. Am. A*, 1999, v. 16, 2914–2916.

33. Mair A., Vaziri A., Weihs G., Zeilinger A. Entanglement of the orbital angular momentum states of photons. *Nature*, 2001, v. 412, 313–316.

34. Franke-Arnold S., Barnett S. M., Padgett M. J., Allen L. Two-photon entanglement of orbital angular momentum states. *Phys. Rev. A*, 2002, v. 65, 033823-1.

35. Hestenes D. Oersted Medal Lecture 2002: Reforming the Mathematical Language of Physics. *Am. J. Phys.*, v. 71, 104.

36. Muralidhar K. The Spin Bivector and Zeropoint Energy in Geometric Algebra. *Adv. Studies Theor. Phys.*, 2012, v. 6, 675–686.

37. Muralidhar K. Complex Vector Formalism of Harmonic Oscillator in Geometric Algebra: Particle Mass, Spin and Dynamics in Complex Vector Space. *Found. Phys.*, 2014, v. 44, 265–295.

38. Muralidhar K. Algebra of Complex vectors and Applications in electromagnetic Theory and Quantum Mechanics. *Mathematics*, 2015, v. 3, 781–842.

39. Roychoudhuri C., Kracklauer A. F., Creath K. The Nature of Light. What is a Photon? C. R. C. Press, Taylor and Francis Group, Boca Raton, FL, 2008.

40. Doran C., Lasenby A. Geometric Algebra for Physicists. Cambridge University Press, Cambridge, 2003.

41. Kobe D. H. A relativistic Schrödinger-like Equation for a Photon and Its second Quantization. *Found. Phys.*, 1999, v. 29, 1203–1231.

42. Nienhuis G., Allen L. Paraxial wave optics and harmonic oscillators. *Phys. Rev. A*, 1993, v. 48, 656–665.

Gravitational Waves from a Sinusoidially Varying Spherical Distribution of Mass

L. F. Obagboye[1], S. X. K. Howusu[1], E. N. Chifu[2] and E. J. N. Omaghali[3]

[1]Theoretical Physics Programme, National Mathematical Centre Abuja, Nigeria
E-mail: sxkhowusu@yahoo.com
[2]Physics Department, Federal University Dutse, Nigeria
E-mail: ebenechifu@yahoo.com
[3]Physics Department, University of Jos, Nigeria

A theory is developed for the study of spherical gravitational waves by constructing a Generalized Gravitational Field Equation from Newton's gravitational field equation. The Euclidean Laplacian ∇^2 is replaced with the Riemannian Laplacian ∇_R^2. A general gravitational field equation is obtained which resolves the incompleteness in Newton's gravitational field equation. The general gravitational field equation reduces to the pure Newtonian gravitational field equation in the limit of c^0 as required by the Principle of Equivalence of Physics. It also contains post Newton correction terms of orders of c^{-2} and all degrees of nonlinearity in the gravitational scalar potential and its derivatives. Considering a sinusoidally varying homogeneous spherical distribution of mass in the frame work of the obtained general gravitational field equation, gravitational waves are predicted with phase velocity equivalent to the speed of light in vacuo.

1 Introduction

According to General Relativity Theory, gravitational waves are oscillations of spacetime or small distortions of spacetime geometry, or ripples of spacetime curvature which propagate in the time through space as waves. Gravitational waves are produced mainly by extremely massive binary stellar objects, such as binary neutron stars or binary black holes. Though gravitational waves can be produced by all mass interactions, the amplitude of these waves is far too small to be detected. Normal solar systems produce gravitational waves when their planets orbit their primary, but again, these are incredibly tiny ripples. Even a binary black hole — which produces the most powerful gravitational waves we can imagine — requires measurements of distances of about 1/1000 of the diameter of a proton [1].

The search for gravitational waves has been the centre of current research in Astronomy and Cosmology. Higher precision and more sensitive detectors have been developed over the years. Experiments on gravitational waves started with Weber's experiments on gravitational antennae; in which he registered weak signals [2]. He concluded that some processes at the centre of the Galaxy were the origin of the detected signals. Other attempts were made in detecting gravitational waves such as [3-6]. The most recent experimental attempt by Abbott *et al.* in 2015 [7] claims that two detectors of the Laser Interferometer Gravitational-Wave Observatory simultaneously observed a transient gravitational-wave signal.

Much theoretical work has also been done to either proof or disproof the existence of gravitational waves. In a nutshell, theoretical studies of gravitational waves can be classified into three main groups [2]:

- Research targeted at giving an invariant definition for gravitational waves. These include Pirani [8], Bondi [9], and others.
- Searching for solutions to Einstein gravitational field equations by proceeding from physical considerations to describe gravitational radiations. These include studies by Einstein and Rosen [10], Petrov [11], Chifu and Taura [1] and others.
- Studying gravitational inertial waves, covariant with respect of transformations of spatial coordinates and also invariant with respect of transformations of time [12].

This research article falls in the second group. The so-called "Great Metric Tensor" [13-14] is used to deduce a general gravitational wave equation; which is later applied to a sinusoidally varying mass for a homogeneous spherical distribution of mass.

2 The general spherical gravitational field equation

Newton's gravitational field equation is given by

$$\nabla^2 f(\underline{r}, t) = 4\pi G \rho_0(\underline{r}, t) \tag{1}$$

where, ρ_0 is the density of proper mass in a distribution or system, ∇^2 is the pure Euclidean Laplacian, G is the universal gravitational constant and f is the gravitational scalar potential.

The incompleteness of equation (1) are as follows:

1. The density of proper mass (source of gravitational field) in equation (1) can vary with coordinate time and the Euclidean Laplacian cannot account for this possible variation.

2. Time variation of proper mass should result in the radiation of energy possibly in the form of gravitational waves or radiation that can propagate in space-time with or without gravitational field.
3. Newton's gravitational intensity vector $\underline{g}$ is given by

$$\underline{g} = -\underline{\nabla} f \tag{2}$$

where $\underline{\nabla}$ is the Euclidean gradient operator.

The Euclidean operator in equation (2) above has no variation with time and hence will not be sufficient for the complete description of gravitational intensity vector of time dependent gravitational fields.

From the foregoing it becomes necessary to seek a general gravitational field equation which will be sufficient for the description of all gravitational fields. Howusu in 2009 [13] proposed that a general gravitational field equation based on Riemannian coordinate geometry may be obtained by replacing the Euclidean Laplacian with Riemannian Laplacian to obtain

$$\nabla_R^2 f(\underline{r}, t) = 4\pi G \rho_0(\underline{r}, t) \tag{3}$$

where ∇_R^2 is the Riemannian Laplacian based on the great metric tensor for all possible gravitational fields. The gravitational intensity (acceleration due to gravity) for all possible gravitational fields can also be defined in terms of the Riemannian gradient operator ∇_R. The most general form of the Riemannian Laplacian is given as

$$\nabla_R^2 = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\mu} \left(\sqrt{g} g^{\mu\nu} \frac{\partial}{\partial x^\nu} \right) \tag{4}$$

where $g^{\mu\nu}$ is the contravariant metric tensor. Thus, for any function $f(\underline{r}, t)$ we can write

$$\begin{aligned} \nabla_R^2 f(\underline{r}, t) = & \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i} \left(\sqrt{g} g^{ij} \frac{\partial}{\partial x^j} \right) f(\underline{r}, t) \\ & + \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^0} \left(\sqrt{g} g^{00} \frac{\partial}{\partial x^0} \right) f(\underline{r}, t). \end{aligned} \tag{5}$$

Using Einstein's coordinates with $x^0 = ct$, equation (5) can be written explicitly as

$$\begin{aligned} \nabla_R^2 f(\underline{r}, t) = & \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i} \left(\sqrt{g} g^{ij} \frac{\partial}{\partial x^j} \right) f(\underline{r}, t) \\ & + \frac{1}{c^2 \sqrt{g}} \frac{\partial}{\partial t} \left(\sqrt{g} g^{00} \frac{\partial}{\partial t} \right) f(\underline{r}, t). \end{aligned} \tag{6}$$

Hence equation (3) can be written more explicitly as

$$\begin{aligned} 4\pi G \rho_0(\underline{r}, t) = & \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i} \left(\sqrt{g} g^{ij} \frac{\partial}{\partial x^j} \right) f(\underline{r}, t) \\ & + \frac{1}{c^2 \sqrt{g}} \frac{\partial}{\partial t} \left(\sqrt{g} g^{00} \frac{\partial}{\partial t} \right) f(\underline{r}, t). \end{aligned} \tag{7}$$

Equation (7) is the general field equation which resolves the incompleteness of Newton's gravitational field equation. Remarkably, the general gravitational field equation reduces to the pure Newton's gravitational field equation in the limit of c^0 (as required by the Principle of Equivalence of Physics). It may also be noted that the gravitational field equation contains post Newton correction terms of orders of c^{-2} and all degrees of nonlinearity in the gravitational scalar potential and its derivatives.

The Great Metric Tensor for all spherical gravitational fields in spherical polar coordinates (r, θ, ϕ, x^0) is given as [13-14]:

$$g_{11}(r, \theta, \phi, x^0) = \left(1 + \frac{2}{c^2} f(r, \theta, \phi, x^0) \right)^{-1}, \tag{8}$$

$$g_{22}(r, \theta, \phi, x^0) = r^2, \tag{9}$$

$$g_{33}(r, \theta, \phi, x^0) = r^2 \sin^2 \theta, \tag{10}$$

$$g_{00}(r, \theta, \phi, x^0) = -\left(1 + \frac{2}{c^2} f(r, \theta, \phi, x^0) \right) \tag{11}$$

where f is the gravitational scalar potential. From equation (8) to (11) it can be deduced that

$$\sqrt{g} = r^2 \sin \theta. \tag{12}$$

Equation (7) can thus be written as:

$$\begin{aligned} 4\pi G \rho_0(r, t) = & \frac{1}{r^2} \frac{\partial}{\partial r} \left[\left(1 + \frac{2}{c^2} f \right) r^2 \right] f \\ & + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) f \\ & + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} f \\ & - \frac{1}{c^2} \frac{\partial}{\partial t} \left[\left(1 + \frac{2}{c^2} f \right)^{-1} \frac{\partial}{\partial t} f \right]. \end{aligned} \tag{13}$$

Equation (13) is the general spherical gravitational field equation interms of the great metric tensor. The following important facts can be drawn from equation (13):

1. It contains the $\left(1 + \frac{2}{c^2} f\right)$ term which is not found in Newton's gravitational field equation. The consequence of this is that it predicts correction terms to the gravitational field of all massive spherical bodies.
2. The time component of this equation predicts the existence of gravitational waves with velocity which is equal to the speed of light in vacuo.

3 Special case: sinusoidally varying homogenous spherical distribution of mass

Now, consider a sinusoidally varying homogenous spherical distribution of mass. In this case, the mass varies in such a

way that f is independent of the polar angle θ and the azimuthal angle ϕ, [15] such that equation (13) reduces to

$$4\pi G\rho_0(r,t) = \frac{1}{r^2}\frac{\partial}{\partial r}\left[\left(1+\frac{2}{c^2}f\right)r^2\frac{\partial}{\partial r}\right]f - \frac{1}{c^2}\frac{\partial}{\partial t}\left[\left(1+\frac{2}{c^2}f\right)^{-1}\frac{\partial}{\partial t}f\right]. \quad (14)$$

Linearizing equation (14) we obtain:

$$f'' + \frac{2}{r}f' - \frac{1}{c^2}\ddot{f} = 4\pi G\rho_0. \quad (15)$$

Suppose we have a dipole antenna which consists of two spherical bodies where electrons are driven by an oscillator [1]; then the movement of the electric charges driven by the oscillator is equivalent to an exponential factor. We therefore modify equation (15) in such a way that the proper mass density varies sinusoidally within a homogeneous spherical mass distribution such that:

$$f'' + \frac{2}{r}f' - \frac{1}{c^2}\ddot{f} = 4\pi G\rho_e\, e^{i\omega t}. \quad (16)$$

In order to solve equation (16) we seek a solution such that

$$f(r,t) = R(r)\, e^{i\omega t} \quad (17)$$

where R is the radius of the spherical mass distribution. Equation (15) will thus become

$$R''(r) + \frac{2}{r}R'(r) + \frac{1}{c^2}\omega^2 R(r) = 4\pi G\rho_e. \quad (18)$$

Let

$$R(r) = \frac{1}{r}F(r),$$

then

$$R' = -\frac{1}{r^2}F(r),$$

and

$$R''(r) = \frac{1}{r}F''(r) - \frac{2}{r^2}F'(r) + \frac{2}{r^3}F(r).$$

It therefore follows that equation (18) becomes

$$\frac{1}{r}F''(r) + \frac{\omega^2}{c^2 r}F(r) = 4\pi G\rho_e. \quad (19)$$

Hence, the interior field equation for this distribution of mass is given as

$$\frac{1}{r}F''(r) + \frac{\omega^2}{c^2 r}F(r) = 4\pi G\rho_e\,;\ r < R \quad (20)$$

and the corresponding exterior field equation as:

$$\frac{1}{r}F''(r) + \frac{\omega^2}{c^2 r}F(r) = 0\,;\ r > R. \quad (21)$$

Equation (21) is a simple harmonic function which can have three solutions viz:

$$F(r) = Be^{ikr}, \quad (22)$$

$$F(r) = D\cos(kr), \quad (23)$$

and

$$F(r) = E\sin(kr). \quad (24)$$

Taking the first and second derivatives of equation (22) we have

$$F'(r) = ikB\, e^{ikr}$$

and

$$F''(r) = -k^2 B\, e^{ikr},$$

which can be substituted into (21) to yield

$$-k^2 B\, e^{ikr} + \frac{\omega^2}{c^2}B\, e^{ikr} = 0, \quad (25)$$

hence

$$k = \pm\frac{\omega}{c}. \quad (26)$$

We thus state the complimentary solution as

$$F_c^-(r) = E\sin\left(\frac{\omega}{c}r\right);\ r > R \quad (27)$$

$$F_c^+(r) = D\cos\left(\frac{\omega}{c}r\right);\ r < R. \quad (28)$$

The particular solution for the interior field equation is given by

$$\frac{1}{r}F''(r) + \frac{\omega^2}{c^2 r}F(r) = 4\pi G\rho_e; r < R. \quad (29)$$

Let $F(r) = Ar$, then $F'(r) = A$ and $F''(r) = 0$ and equation(29) yields

$$A = \frac{4\pi G c^2\rho_e}{\omega^2}, \quad (30)$$

and hence

$$F_p^-(r) = \frac{4\pi G c^2\rho_e}{\omega^2}r. \quad (31)$$

Equation (31) is thus the particular solution for the exterior field equation. The general solution for the exterior field is then given as

$$R^+(r) = \frac{D}{r}\cos\left(\frac{\omega}{c}r\right)\frac{4\pi G c^2\rho_e}{\omega^2}. \quad (32)$$

Equation (17) can thus be fully expressed as

$$f^+(r,t) = \frac{D}{r}\cos\left(\frac{\omega}{c}r\right)\cos(\omega t) + \frac{iD}{r}\cos\left(\frac{\omega}{c}r\right)\sin(\omega t) \quad (33)$$

with independent solutions

$$f^+(r,t) = \frac{D}{r}\cos\left(\frac{\omega}{c}r\right)\cos(\omega t) \quad (34)$$

and

$$f^{+}(r,t) = \frac{D}{r}\cos\left(\frac{\omega}{c}r\right)\sin(\omega t). \tag{35}$$

The two solutions (34) and (35) can be combined to yield

$$f^{+}(r,t) = \frac{1}{2}\frac{D}{r}\left[\cos\left(\omega\left(\frac{r}{c}+t\right)\right)+\cos\left(\omega\left(\frac{r}{c}-t\right)\right)\right]. \tag{36}$$

From equation (36) it is clear that the phase of the wave equation ϕ is given by

$$\phi = \frac{\omega r}{c} \pm \omega t, \tag{37}$$

hence

$$\frac{dr}{dt} = c. \tag{38}$$

4 Concluding remarks

In this paper we have shown [equation (36)] that in the limit of linear terms, the general gravitational field equation predicts gravitational waves with phase velocity which is equal to the speed of light in empty space. These waves will not vary with any angle, hence they will move along radial lines from inside the sphere outwards(radial waves). A sinusoidally varying mass thus radiates spherical gravitational waves. The obtained results gives similarlar predictions as in [1, 16] in the limit c^{-2} though in the limit c^{0} [16] predicts gravitational waves with imaginary phase.

Submitted on May 27, 2016 / Accepted on June 2, 2016

References

1. Chifu E.N. and Taura L.S. *Progress in Physics*, 2013, v. 9, issue 3, 7.
2. Borissova L. *Progress in Physics*, 2005, v. 1, issue 2, 30.
3. Braginsky V.B., Manukin A.B., Popov E.J., Rudenko V.N. *Phys. Rev. Ser. A*, 1973, v. 45, 271.
4. Douglas D.N., Gram R.Q., Tyson J.A., Lee R.W. *Phys. Rev. Lett.*, 1975, v. 35, 480.
5. Levine J.L., Garwin R.L. *Phys. Rev. Lett.*, 1973, v. 31, 173.
6. Tyson J.A., Macleannan C.G., Lanzorotti L.J. *Phys. Rev. Letts*, 1973, v. 30, 1006.
7. Abbott B.P. et al. *Phys. Rev. Lett*, 2015, v. 116, 061102.
8. Bondi H., Pirani F., Robinson J. *J. Proc. Roy. Soc. A*, 1959, v. 251, 519.
9. Bondi H. *Nature*, 1957, v. 179, 1072.
10. Rosen N. *Res. Council Israel*, 1954, v. 4, 328.
11. Petrov A.Z. Einstein Spaces. Pergamon, london, 1969.
12. Borissova L. *The Abraham Zelmanov Journal*, 2010, v. 3, 26.
13. Howusu S.X.K. The Metric Tensors for Gravitational Fields Physics and the Mathematical Principles of Riemannian Theoretical Physics. Jos University Press, 2009.
14. Howusu S.X.K. Riemannian Revolutions in Physics and Mathematics. Jos University Press, 2013.
15. Chifu E.N. and Howusu S.X.K. *Physics Essays*, 2009, v. 22(1), 73.
16. Chifu E.N. and Howusu S.X.K. *Progress in Physics*, 2009, v. 5, issue 3, 45.

Gravitational Shielding as Viewed in the Planck Vacuum Theory

William C. Daywitt

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

This paper argues that gravitational shielding does not exist, because gravitational waves travel within the vacuum state rather than free space.

1 Introduction

The concept of gravitational shielding has been around for a long time and it would incorrectly assert, for example, that when the earth lines up between the moon and the sun, the moon-sun gravitational attraction is reduced. The fact that this shielding (by the earth in this case) does not occur is one of the great mysteries in the history of physics. The theory of the Planck vacuum (PV) state, however, offers an easy explanation for the absence of such shielding.

As a counter example to the gravitational force, consider the free space Coulomb force (e^2/r^2) between two charges e separated by the distance r. If a shield of any type whatsoever is placed between the charges, the resulting force is changed dramatically. Indeed, if a large enough grounded screen were inserted between the charges, the force would vanish entirely.

2 Newton Force

Now consider the gravitational force between two free space masses m in the center-of-mass (CoM) coordinate frame defined by $m\mathbf{r}_1 + m\mathbf{r}_2 = 0$ (with $-\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$):

$$F_{gr}(r) = -\frac{m^2G}{(2r)^2} = -\frac{m^2c^4}{4r^2(c^4/G)} = -\frac{(mc^2/2r)^2}{(c^4/G)} \tag{1}$$

$$= -\frac{(mc^2/2r)^2}{(m_*c^2/r_*)} = -n_{2r}^2\frac{m_*c^2}{r_*} \tag{2}$$

where the n-ratio

$$n_{2r} = \frac{mc^2/2r}{m_*c^2/r_*} < 1 \tag{3}$$

is the normalized force either mass m exerts on the PV at the position of the opposite mass, where the masses are centered at $\pm\mathbf{r}$ from the origin of the CoM coordinates. The normalization force m_*c^2/r_* is the maximum force the PV can sustain before breaking down. This force also normalizes the Einstein field equation [1, eqn.15].

The three ratios in (1) are the force equation expressed in terms of Newton's secondary constant G, an experimental constant that makes (1) agree with the experimental data. As such, however, G hides a significant amount of physics. The substitution $c^4/G = m_*c^2/r_*$ [1, eqn.5] replaces G by a combination of primary (fundamental) constants that lead to (2) and the following nonrelativistic explanation of the gravitational force.

The gravitational field $g(r)$ of either mass can be defined in the usual manner and yields

$$g(r) = \frac{F_{gr}(r)}{m} = -\frac{c^2n_{2r}}{2r} \tag{4}$$

which is again centered at the radii $\pm r$ from the CoM origin, where r is the coordinate radius common to both free space and its underlying PV state.

From (2) and (4) it is easy to carry these calculations a step further. Newton's second law applied to either mass gives the acceleration

$$\ddot{r} = \frac{d\dot{r}}{dt} = \dot{r}\frac{d\dot{r}}{dr} = -\frac{c^2n_{2r}}{2r} = -\frac{c^2\cdot mc^2}{4\cdot m_*c^2/r_*}\frac{1}{r^2} \tag{5}$$

or

$$\dot{r}d\dot{r} = -\frac{c^2\cdot mc^2}{4\cdot m_*c^2/r_*}\frac{dr}{r^2} \tag{6}$$

of the masses. Integrating both sides of (6) from r_0 to r leads to

$$\frac{\dot{r}^2 - \dot{r}_0^2}{2} = \frac{c^2n_{2r0}}{2}\left(\frac{r_0}{r} - 1\right) \tag{7}$$

where $\dot{r}_0$ is the velocity of either mass at $r = r_0$, and $r \le r_0$. Without changing the final conclusions, it is convenient to set $\dot{r}_0 = 0$ — i.e., to assume that the masses are released from rest at $\pm r_0$. Then (7) yields their relative velocities toward the origin

$$\frac{\dot{r}}{c} = -\left[n_{2r0}\left(\frac{r_0}{r} - 1\right)\right]^{1/2} \tag{8}$$

where

$$n_{2r0} = \frac{mc^2/2r_0}{m_*c^2/r_*} \tag{9}$$

and $[\cdots]$ in (8) is the normalized force either mass exerts on the PV at the position of the other mass.

3 Conclusions

Three important observations are evident from the previous calculations: equations (2), (4), and (8) are all expressed in terms of PV parameters, implying that the vacuum state mediates the dynamics of the gravitational force between free space masses. A corollary to this conclusion is that gravitational waves, the carrier of the gravitational force, do not propagate in free space — they propagate within the degenerate PV state. Thus free-space gravitational shielding does not change the gravitational force between free space masses.

Finally, the fact that the PV is a degenerate state implies that the Planck particles making up the PV quasi-continuum cannot execute macroscopic (as opposed to microscopic) motions. Thus the gravitational waves that propagate through the PV state must be percussion-like waves, similar to the waves traveling on the surface of a kettle drum.

Submitted on May 26, 2016 / Accepted on June 5, 2016

References

1. 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, no. 1-1, 22. See also www.planckvacuum.com.

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 acessed 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